

École doctorale des Sciences de la Vie et de la Santé

UNIVERSITÉ D'AIX MARSEILLE I

THÈSE

*Présentée et soutenue publiquement le 19 Novembre 2009 par*

**Clémence Roger**

En vue de l'obtention du grade de

**DOCTEUR DE L'UNIVERSITÉ**

*Spécialité : Neurosciences*

**Supervision de l'action et optimisation des comportements :**

**Études électrophysiologiques et IRMf**

Membres du Jury

Pr. Franck VIDAL	Président	LNC, UMR 6155, CNRS, Marseille, France
Dr. Emmanuel PROCYK	Rapporteur	U846, INSERM, Lyon, France
Dr. Simon THORPE	Rapporteur	CerCo, UMR 5549, CNRS, Toulouse, France
Dr. Etienne KOEHLIN	Examineur	U960, INSERM, Paris, France
Pr. Markus ULLSPERGER	Examineur	MPI for Neurological Research, Cologne, Allemagne
Dr. Borís BURLE	Directeur	LNC, UMR 6155, CNRS, Marseille, France



# Table des matières

<b>I</b>	<b>Introduction</b>	<b>9</b>
<b>1</b>	<b>Le contrôle cognitif</b>	<b>11</b>
1.1	Mise en évidence du contrôle . . . . .	11
1.1.1	Les tâches de temps de réaction . . . . .	11
1.1.2	Les ajustements post-erreurs . . . . .	12
1.1.3	Effets de compatibilité et effets séquentiels . . . . .	12
1.1.4	Électromyographie (EMG) et décomposition du TR . . . . .	13
1.2	Structures impliquées . . . . .	16
1.2.1	Cortex Cingulaire Antérieur et Cortex Préfrontal chez l'Homme . . . . .	16
1.2.2	Ganglions de la Base . . . . .	20
1.3	Indices EEG . . . . .	23
1.3.1	L'onde d'erreur ( $N_E$ ou ERN) . . . . .	23
1.3.2	La négativité liée au feedback (FRN) . . . . .	24
1.3.3	La $N_E$ de l'observateur . . . . .	25
1.3.4	La $N_E$ , la FRN et la $N_E$ de l'observateur . . . . .	26
1.3.5	La Positivité d'Erreur ( $P_E$ ) . . . . .	26
1.3.6	Activation et Inhibition des cortex moteurs . . . . .	27
1.3.7	La N-40 . . . . .	29
1.3.8	La $N2$ . . . . .	29
1.4	La spécificité de la $N_E$ . . . . .	30
1.4.1	La $N_E$ est-elle spécifique aux erreurs? . . . . .	31
1.4.2	La taille de la $N_E$ varie-t-elle en fonction du nombre d'éventualités? . . . . .	32
1.4.3	La $N_E$ est-elle sensible à la compatibilité? . . . . .	34
1.4.4	La taille de la $N_E$ peut-elle présager de la performance à l'essai suivant? . . . . .	35

1.4.5	Conclusion . . . . .	36
1.5	Pathologies et Contrôle . . . . .	36
1.5.1	Effets dûs au développement normal . . . . .	37
1.5.2	Hypofonctionnement des processus de contrôle . . . . .	38
1.5.3	Hyperfonctionnement des processus de contrôle . . . . .	39
1.5.4	Les effets des manipulations expérimentales . . . . .	39
1.5.5	Bilan . . . . .	40
<b>2</b>	<b>Modèles</b>	<b>43</b>
2.1	Introduction aux Réseaux de Neurones . . . . .	43
2.2	Modèles modélisant l'activité du CCA . . . . .	46
2.2.1	Hypothèse de la Mismatch . . . . .	46
2.2.2	Hypothèse Émotionnelle . . . . .	47
2.2.3	Le modèle du conflit . . . . .	48
2.2.4	Modèle WEAVER ++ (Roelofs, 2003) . . . . .	60
2.3	Modèles Ganglions de la Base . . . . .	67
2.3.1	Modèle d'Apprentissage par Renforcement . . . . .	67
2.3.2	Modèle de Ganglions de la Base . . . . .	72
2.4	Conclusion sur les modèles . . . . .	76
<b>II</b>	<b>Partie Expérimentale</b>	<b>79</b>
<b>3</b>	<b>Conflit et Ne</b>	<b>81</b>
3.1	Introduction . . . . .	81
3.2	Methods . . . . .	85
3.2.1	Simulation study . . . . .	85
3.2.2	Experimental study . . . . .	86
3.2.3	Temporal overlap estimation . . . . .	87
3.2.4	Single trial analysis . . . . .	88
3.3	Results . . . . .	91
3.3.1	Behavioral data . . . . .	91
3.3.2	Comparison between conflict and $N_E$ timing as a function of $\Delta$ . . . . .	91
3.3.3	Conflict and the dynamics of incorrect and correct response activation . . .	95



3.4	Discussion . . . . .	100
3.4.1	$N_E$ and conflict . . . . .	100
3.4.2	Response co-activation, conflict and behavioral interference . . . . .	102
3.4.3	Relations to other models of ACC/ $N_E$ function . . . . .	103
3.4.4	Online control of control? . . . . .	105
3.5	APPENDIX A . . . . .	106
3.6	APPENDIX B . . . . .	107
<b>4</b>	<b>ICA and Ne</b>	<b>111</b>
4.1	Introduction . . . . .	111
4.2	Materials and Methods . . . . .	114
4.2.1	Subjects . . . . .	114
4.2.2	Task, Recordings and EEG Data preprocessing . . . . .	114
4.2.3	EEG Data analysis . . . . .	115
4.2.4	Blind Source Separation (BSS) : General Principle . . . . .	115
4.2.5	ICA in the present study . . . . .	116
4.2.6	Source Localization . . . . .	118
4.3	Results . . . . .	118
4.3.1	Behavioral Data . . . . .	118
4.3.2	EEG Data . . . . .	118
4.3.3	ICA Data . . . . .	120
4.3.4	Source Localization . . . . .	121
4.3.5	Lateralization of the $N_C$ on correct trials? . . . . .	123
4.4	Discussion . . . . .	125
<b>5</b>	<b>Sélection de la réponse</b>	<b>129</b>
5.1	Introduction . . . . .	129
5.2	Expérience 1 . . . . .	132
5.2.1	Introduction . . . . .	132
5.2.2	Matériel - Méthode . . . . .	132
5.2.3	Résultats Comportementaux . . . . .	136
5.2.4	Résultats Électrophysiologiques . . . . .	137
5.2.5	Discussion de l'expérience 1 : . . . . .	142
5.3	Expérience 2 . . . . .	143

5.3.1	Introduction . . . . .	143
5.3.2	Matériel - Méthode . . . . .	143
5.3.3	Résultats comportementaux . . . . .	145
5.3.4	Résultats Électrophysiologiques . . . . .	146
5.3.5	Discussion de l'expérience 2 : . . . . .	150
5.4	Discussion des 2 expériences . . . . .	150
5.4.1	Rôle des cortex moteurs primaires . . . . .	152
5.4.2	Le rôle de la N-40 . . . . .	154
5.4.3	La négativité prémotrice ipsilatérale à la réponse . . . . .	155
5.5	Conclusion . . . . .	155
<b>6</b>	<b>Attention et processus de contrôle</b>	<b>157</b>
6.1	Introduction . . . . .	157
6.2	Méthode . . . . .	160
6.2.1	Les sujets . . . . .	160
6.2.2	Les stimuli . . . . .	161
6.2.3	Dessin expérimental . . . . .	161
6.2.4	Acquisition des données . . . . .	162
6.2.5	Traitement des données . . . . .	163
6.3	Résultats . . . . .	164
6.3.1	Résultats Comportementaux . . . . .	164
6.3.2	Résultats IRMf . . . . .	165
6.4	Discussion . . . . .	169
<b>III</b>	<b>Discussion Générale</b>	<b>173</b>
<b>7</b>	<b>Discussion Générale</b>	<b>175</b>
7.1	Nature de la $N_E$ . . . . .	175
7.1.1	$N_E$ et erreurs . . . . .	175
7.1.2	$N_E$ et Conflit . . . . .	175
7.1.3	$N_E$ : un mécanisme général? . . . . .	176
7.2	Lien avec les autres ondes . . . . .	178
7.2.1	$N_E$ et FRN . . . . .	178

---

7.2.2	$N_E$ et $N_2$ . . . . .	178
7.3	Ajustements Comportementaux . . . . .	181
7.3.1	Ralentissement post-erreur . . . . .	181
7.3.2	Accélération pré-erreur . . . . .	182
7.3.3	Adaptation à la compatibilité . . . . .	182
7.4	Conclusions . . . . .	183
	Bibliographie . . . . .	184
	Table des figures . . . . .	197
	Appendix : papiers non directement liés à la thèse . . . . .	201
<b>A</b>	<b>Autres papiers</b>	<b>203</b>



## Première partie

### Introduction



# Chapitre 1

## Le contrôle cognitif

Une caractéristique importante des processus de traitement de l'information chez l'Homme est leur aptitude à s'ajuster à un environnement changeant. Cette flexibilité semble être possible grâce à des mécanismes de contrôle qui organisent et ajustent les processus de traitement de l'information en fonction du contexte, c'est à dire que l'organisation du traitement ne dépend pas seulement de la nature des stimuli à traiter (ou des réponses à fournir) mais aussi de processus descendants (*top-down*) qui configurent de manière adaptée les opérations requises. Un tel contrôle cognitif a été mis en évidence dans une grande variété de tâches.

### 1.1 Mise en évidence du contrôle

#### 1.1.1 Les tâches de temps de réaction

Dans les tâches de temps de réaction (tâche de TR), on demande aux sujets d'identifier un stimulus et de répondre le plus vite possible en évitant de faire des erreurs. Dans les “tâches de TR simple”, un seul stimulus est présenté, et une seule réponse est possible. Les sujets doivent appuyer sur un bouton réponse dès l'apparition du stimulus. Dans les “tâches de TR de choix” plusieurs réponses sont possibles; le sujet doit choisir sa réponse en fonction de la nature du stimulus. Le sujet doit identifier le stimulus, trouver la réponse qui est associée à ce stimulus en fonction de l'association stimulus-réponse apprise au préalable puis choisir et d'exécuter le programme moteur correspondant à cette réponse.

La figure 1.1 illustre une tâche de TR de choix à deux éventualités dans laquelle on demande aux sujets de répondre en appuyant sur le bouton de droite lorsque le stimulus est vert, et sur

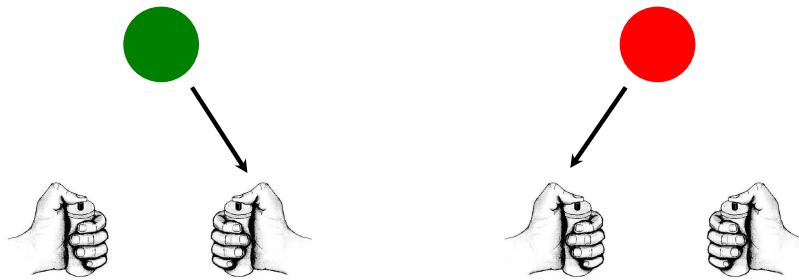


FIGURE 1.1 – **Exemple d’une tâche de TR de choix à deux éventualités.** La consigne est d’identifier la couleur du stimulus et de répondre en appuyant sur le bouton poussoir placé dans la main droite si le stimulus est vert, et sur le bouton poussoir placé dans la main gauche si le stimulus est rouge.

le bouton de gauche lorsque le stimulus est rouge. Le TR correspond à l’intervalle de temps qui sépare la présentation du stimulus de l’enregistrement de la réponse mécanique.

### 1.1.2 Les ajustements post-erreurs

Rabbitt (1966) ont montré que, dans les tâches de TR, les sujets ralentissent dans les essais suivant une erreur. Ce “ralentissement post-erreur”, illustré dans la figure 1.2, suggère qu’il existe bien un système chargé du contrôle de l’exécution qui opère d’un essai à l’autre. La probabilité de commettre une erreur (à l’essai  $n$ ) est plus faible lorsque, à l’essai précédent ( $n-1$ ) une erreur a été commise plutôt qu’une réponse correcte (Laming, 1979).

### 1.1.3 Effets de compatibilité et effets séquentiels

La manipulation de la compatibilité Stimulus-Réponse (S-R) dans les tâches de TR est utilisée pour mettre en évidence les processus de contrôle. La tâche de Simon est l’une des tâches manipulant la compatibilité S-R (Craft et Simon, 1970). Dans cette tâche, les sujets ont pour consigne d’identifier la couleur d’un stimulus et de répondre en fonction de celle-ci. L’information sur la position du stimulus n’est pas pertinente pour la tâche. Les essais compatibles sont ceux pour lesquels le stimulus est présenté du côté de la réponse, alors que les essais incompatibles sont ceux pour lesquels le stimulus est présenté du côté opposé à la réponse (figure 1.3). Les TR sont plus courts en réponse aux stimuli compatibles qu’aux stimuli incompatibles. La différence entre les TR dans les essais compatibles et incompatibles est appelé “effet de compatibilité”.

Par ailleurs, lorsque la compatibilité S-R varie d’un essai à l’autre, la taille de l’effet de compatibilité est modulé en fonction de la nature du stimulus  $n-1$  (Gratton *et al.*, 1992; Stürmer *et al.*,



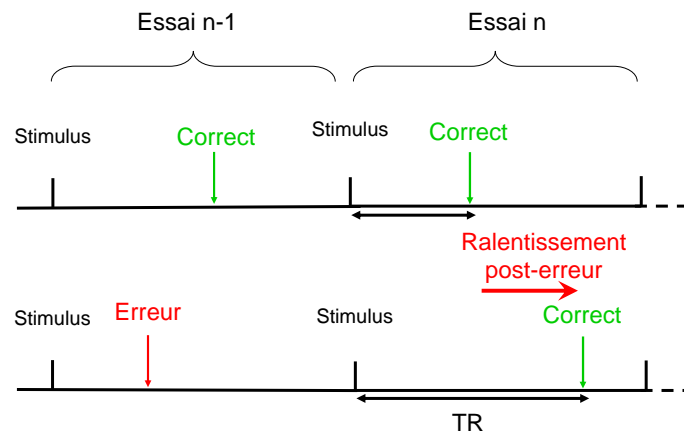


FIGURE 1.2 – **Illustration d’un ajustement post-erreur.** Sur cette figure sont représentés deux séries de deux essais. Le trait horizontal représente le temps. Le premier essai  $n-1$  est suivi de l’essai  $n$ . Le temps de réaction (TR) à l’essai  $n$  dépend de la performance à l’essai  $n-1$ . On observe un allongement du TR à l’essai  $n$  lorsque le sujet a commis une erreur à l’essai  $n-1$ .

2002). L’effet de compatibilité est plus important lorsque l’essai  $n-1$  est compatible que lorsque l’essai  $n-1$  est incompatible (voir figure 1.4). Cet effet est aussi observé sur le taux d’erreur.

#### 1.1.4 Électromyographie (EMG) et décomposition du TR

L’EMG enregistre l’activité électrique produite par les muscles grâce à deux électrodes placées sur la peau au-dessus du muscle d’intérêt. L’activité EMG reflète le recrutement des unités motrices. Sans les tâches de TR, l’enregistrement de l’EMG des muscles impliqués dans la réponse correcte et dans la réponse incorrecte, révèle une partie de la dynamique d’activation des réponses par le système nerveux. Il permet notamment de détecter de petites activations musculaires sous-liminales n’ayant pas déclenché la réponse mais qui révèlent une activation de la réponse incorrecte dans certains essais. De cette façon, on peut distinguer les essais “purs-corrects” dans lesquels la seule bouffée EMG déclenche la réponse du sujet, des essais “corrects avec ébauche d’erreur” dans lesquels une petite bouffée EMG du côté de la réponse correcte a été observée avant la bouffée EMG du côté de la bonne réponse (voir figures 1.5).

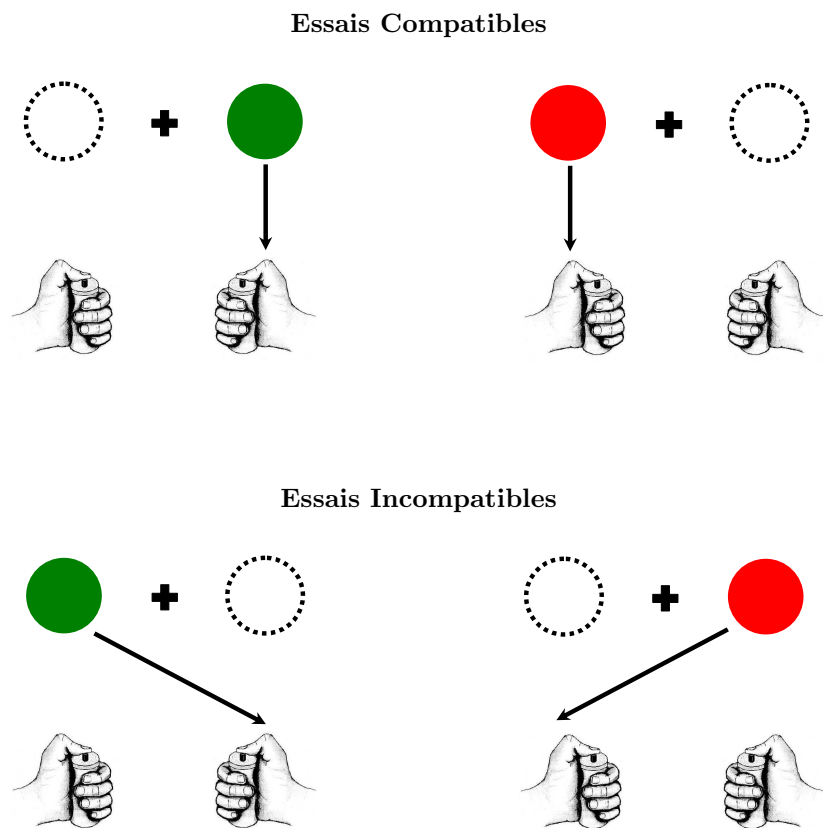


FIGURE 1.3 – **Tâche de Simon manipulant la compatibilité spatiale.** La consigne est d'identifier la couleur du stimulus et de répondre en appuyant sur le bouton poussoir placé dans la main droite si le stimulus est vert, et sur le bouton poussoir placé dans la main gauche si le stimulus est rouge. La position du stimulus est à ignorer. Les essais du haut sont compatibles puisque le stimulus apparaît du côté de la réponse correcte, et ceux du bas sont incompatibles puisque le stimulus apparaît du côté de la réponse incorrecte.

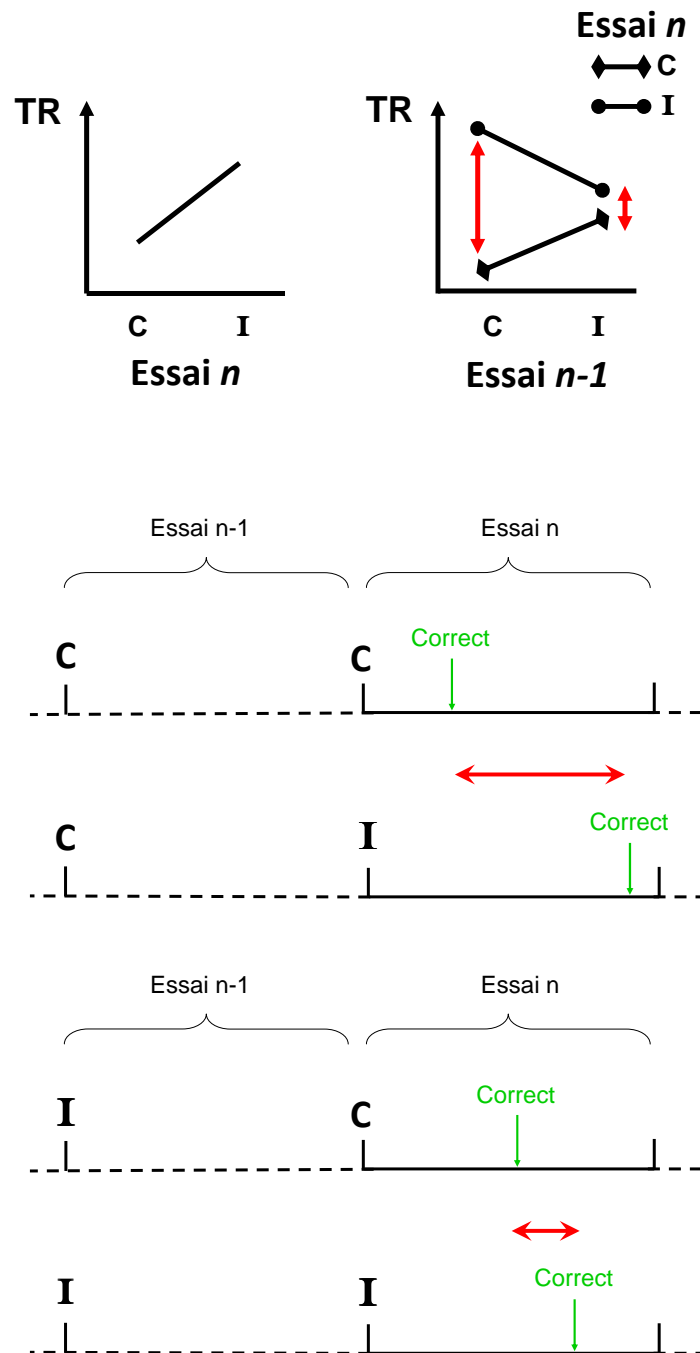


FIGURE 1.4 – Illustration des l'effets séquentiels sur le TR. Sur cette figure sont représentés quatre séries de deux essais. Dans les séries du haut, l'essai *n-1* est compatible (C), dans celles du bas l'essai *n-1* est incompatible (I). La différence de TR entre un essai *n* compatible et un essai *n* incompatible (l'effet de compatibilité, symbolisé par la flèche rouge) est plus grande lorsque l'essai *n-1* est compatible que lorsqu'il est incompatible.

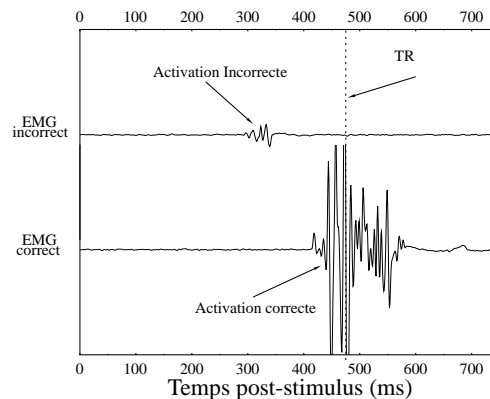


FIGURE 1.5 – **Exemple d’enregistrement de l’activité EMG des muscles impliqués dans un essai correct avec ébauche d’erreur.** Le zéro des temps correspond au moment de présentation du stimulus. L’activité EMG du muscle impliqué dans la réponse incorrecte (tracé du haut) et du muscle impliqué dans la réponse correcte tracé du bas). Le moment de déclenchement de la réponse est indiqué par la ligne pointillée fléchée. On voit apparaître, du côté de la réponse incorrecte, une petite activité EMG que l’on appelle “ébauche d’erreur”.

## 1.2 Structures impliquées dans le contrôle

### 1.2.1 Cortex Cingulaire Antérieur et Cortex Préfrontal chez l’Homme

Le cortex cingulaire antérieur, situé dans la partie interne du lobe frontal (aire 24 de Brodmann, figure 1.6), est une structure qui a été identifiée comme étant impliquée dans les mécanismes de contrôle de l’action.

Pardo *et al.* (1990) ont montré, pour la première fois en utilisant l’imagerie cérébrale, l’implication du cortex cingulaire antérieur (CCA) dans une tâche manipulant la compatibilité stimulus-réponse. En utilisant la tomographie par émission de positons (TEP), les auteurs ont enregistré les variations de débit sanguin cérébral régional chez des sujets sains effectuant la tâche de Stroop (Stroop, 1935). Ils ont comparé les cartes d’activation dans la condition où les stimuli étaient tous compatibles (par exemple, ROUGE BLEU VERT) et dans la condition où les stimuli étaient tous incompatibles (par exemple, VERT ROUGE BLEU). Du fait de l’automatisme du traitement lexical, la lecture des mots est plus rapide que la dénomination des couleurs, elle peut ainsi produire une interférence (dans le cas des essais incompatibles) ou une facilitation (dans le cas des essais compatibles) sur le traitement de la couleur. Il est apparu que le CCA est plus actif dans la condition incompatible que dans la condition compatible. Pour les auteurs, le CCA est impliquée dans la sélection des processus à mettre en œuvre parmi les processus disponibles en fonction des objectifs

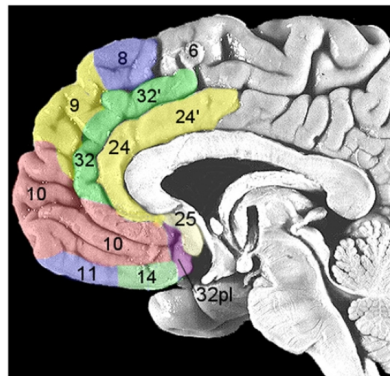


FIGURE 1.6 – **Coupe sagittale médiane du cerveau.** Les nombres font référence aux aires du Brodmann : Aire 6 - Cortex présumé subdivisé en aire présumotrice et aire motrice supplémentaire ; Aire 24 - Cortex cingulaire antérieur, région rostrale ; Aire 24' - Cortex cingulaire antérieur, région caudale (d'après Ridderinkhof *et al.*, 2004b)

donnés. Autrement dit, le rôle du CCA serait d'implémenter la consigne lorsque celle-ci n'est pas "naturelle". En situation incompatible, le rôle du CCA serait de rappeler que la réponse doit être donnée en fonction de la couleur du mot (Pardo *et al.*, 1990).

En utilisant l'imagerie par résonance magnétique fonctionnelle (IRMf), Carter *et al.* (2000) n'ont pas observé de différence d'activité du CCA dans les séries d'essais où le risque d'erreur est important (20% d'essais compatibles, 80% d'essais incompatibles) par rapport aux séries d'essais où le risque d'erreur est faible (80% d'essais compatibles, 20% d'essais incompatibles). En revanche, ils ont noté une augmentation ponctuelle de l'activité du CCA au moment où un fort risque d'erreur survient alors que celui-ci n'est pas attendu (dans les essais incompatibles au sein d'une série d'essais avec faible risque d'erreur). Ces résultats suggèrent que le CCA joue un rôle d'évaluation plutôt qu'un rôle de stratégie. Les résultats de Carter *et al.* (1998) corroborent cette hypothèse puisqu'ils ont montré que le CCA est davantage actif lorsque les sujets commettent une erreur que lorsqu'ils donnent la réponse correcte. Cette interprétation est également en accord avec les résultats de MacDonald *et al.* (2000) dans la tâche de Stroop, qui proposent une complémentarité des rôles du CCA et du cortex préfrontal dorsolatéral (DLPFC<sup>1</sup>) dans les tâches manipulant la compatibilité stimulus-réponse.

Dans le paradigme utilisé par MacDonald *et al.* (2000), chaque essai comprend une période préparatoire et une phase d'exécution. La période préparatoire est initiée par l'apparition d'un indice indiquant au sujet la tâche qu'il doit effectuer (nommer la couleur du mot : tâche COULEUR ; ou bien lire le mot : tâche MOT). La phase d'exécution débute au moment de la présentation

1. DLPFC est l'abréviation anglaise de "dorsolateral prefrontal cortex"

du stimulus. Les auteurs se sont intéressés aux variations d'activité dans le DLPFC et le CCA dans ces deux phases. Pendant la période préparatoire, le DLPFC est plus actif lorsque l'indice informe les sujets qu'ils vont devoir effectuer la tâche COULEUR plutôt que la tâche MOT. Aucune différence d'activation n'est notée dans le CCA. Pendant la phase d'exécution, l'augmentation de l'activité dans le DLPFC est la même quelle que soit la tâche et la nature du stimulus. En revanche, les auteurs ont observé dans la tâche COULEUR que le CCA est plus actif dans les essais incompatibles que dans les essais compatibles. Il semblerait que le DLPFC et le CCA aient des rôles complémentaires. Le DLPFC aurait pour rôle d'implémenter le contrôle cognitif en faisant entrer en jeu des processus de régulation permettant une sélection des comportements appropriés à la tâche, alors que le CCA aurait un rôle d'évaluation du besoin de contrôle cognitif.

Ces hypothèses ont été soutenues par Kerns *et al.* (2004) dans une étude IRMf menée chez des sujets effectuant la version couleur de la tâche de Stroop. Les auteurs se sont intéressés aux variations d'activité du CCA et du cortex frontal en fonction de la compatibilité à l'essai  $n$  et à l'essai  $n-1$ , et de la présence d'ajustements comportementaux. Ils ont montré (1) que le CCA est activé dans les essais précédant des ajustements comportementaux, (2) que le cortex préfrontal est activé dans les essais avec des ajustements comportementaux, et (3) que l'activité du CCA à l'essai  $n-1$  est corrélée à l'activité du cortex préfrontal à l'essai  $n$ . Autrement dit, Kerns *et al.* (2004) ont montré que l'activité du CCA est un bon indicateur de l'activité du cortex préfrontal et de l'ampleur des ajustements comportementaux. L'idée que le CCA engage le contrôle cognitif est précisément en accord avec des hypothèses fondamentales du modèle du conflit que nous exposerons au chapitre 2.

### Le CCA chez le singe :

Les études chez l'animal sont intéressantes puisqu'elles permettent d'avoir une mesure de l'activité cérébrale en effectuant des enregistrements directement dans la structure d'intérêt. L'enregistrement avec des microélectrodes permet de mesurer l'activité unitaire (les potentiels d'action) des neurones distants de  $100\mu m$  de la microélectrode, et permet également d'évaluer les potentiels de champ locaux (local field potential : LFP) relatifs à l'activité d'une population de neurones située à quelques millimètres de l'électrode d'enregistrement.

Ces études chez le singe impliquent un apprentissage de la tâche puis une phase de test. Du fait de la nécessité de passer par un apprentissage, la comparaison intéressante entre la mise en place des mécanismes de contrôle au cours de l'apprentissage et le fonctionnement de ces mécanismes une fois l'apprentissage terminé, est rendue possible.

Procyk *et al.* (2000) ont enregistré les neurones du CCA dans une tâche de découverte et de mémorisation d'une séquence de trois cibles. La sélection de ces cibles dans un ordre précis amenait à la récompense. Dans cette tâche, la phase de recherche de la séquence est à distinguer de la phase de répétition. Pendant la phase de recherche, l'animal identifie et mémorise les cibles amenant à la récompense. Lorsque la séquence correcte est identifiée, l'animal entre en phase de répétition. Il lui suffit de répéter la séquence correcte mémorisée pour obtenir la récompense. Lorsque l'animal a été récompensé dans trois essais consécutifs, une nouvelle séquence de cibles à identifier est présentée à l'animal. Les auteurs se sont intéressés aux neurones du CCA dont l'activité fluctue au cours de la tâche. Ils ont montré que certains neurones sont préférentiellement actifs pendant la phase de recherche et d'autres pendant la phase de répétition. Les neurones actifs pendant la phase de recherche auraient une activité en lien avec la découverte de la récompense. Tandis que les neurones actifs pendant la phase de répétition, auraient une activité en lien avec la prédiction de la récompense. Les résultats obtenus par Amiez *et al.* (2005) vont dans le sens de cette dernière interprétation. Les neurones du CCA, dont l'activité est significativement augmentée au moment d'une erreur, sont également actifs au moment de l'apparition du signal de changement de séquence (dès que l'animal a été récompensé dans trois essais consécutifs). L'augmentation de la fréquence de décharge des neurones du CCA refléterait l'arrivée d'un événement, dont la caractéristique est qu'il n'est pas accompagné d'une récompense. L'interprétation selon laquelle les neurones du CCA auraient leur activité liée à la probabilité d'obtenir la récompense est à rapprocher des travaux de Schultz *et al.* (1997) qui ont observé des modulations de l'activité des neurones du système mésencéphalique dopaminergique en fonction de la possibilité d'anticipation de la récompense. Nous y reviendrons dans le paragraphe 1.2.2.

Dans une autre tâche, Amiez *et al.* (2005) ont montré que les neurones du CCA actifs au moment de l'erreur ont leur activité modulée en fonction la valeur de la récompense perdue. Trois stimuli associés à des valeurs de récompense différentes (1.4 ml, 0.4 ml et 0 ml de liquide renforçant), ont été utilisés. Après quelques essais, les animaux ont associé chaque stimulus avec l'importance de sa récompense. L'activité des neurones présente au moment de l'erreur est d'autant plus grande que la récompense perdue est importante. Pour que ces neurones aient développé cette capacité, il apparaît nécessaire qu'une comparaison entre les différentes récompenses ait été réalisée et qu'un lien entre chaque stimulus et la valeur de sa récompense ait été établi. Le CCA pourrait jouer le rôle d'intégrateur des informations. Pour aller plus loin, Amiez *et al.* (2006) ont utilisé un protocole dans lequel on présente une dizaine de fois la même paire de stimuli. En moyenne, l'un des stimuli est associé à une récompense plus grande que l'autre. À chaque essai, les animaux choisissent l'un

des deux stimuli et sont récompensés. Très rapidement, les animaux identifient quel stimulus est, en moyenne, associé à la plus grande récompense et maintiennent ce choix de réponse jusqu'à la fin de la série. Il y a donc une période de recherche de la récompense optimale, suivie d'une période de répétition. Les auteurs se sont intéressés aux neurones dont l'activité fluctue au cours de la tâche. Ils ont montré que les neurones du CCA intègrent des informations sur les stimuli, les mouvements exécutés ainsi que sur la taille de la récompense. Ceci suggère que le CCA construit des associations stimulus-récompense sur la base des informations encodées. Cette idée a été testée en inactivant ou en lésant le CCA. Les résultats montrent une difficulté des animaux à utiliser les informations sur la récompense pour optimiser leurs choix (Hadland *et al.*, 2003; Amiez *et al.*, 2006; Kennerley *et al.*, 2006).

Plus récemment, Quilodran *et al.* (2008) ont montré la capacité du CCA à encoder des informations sur les résultats positifs ou négatifs en fonction du contexte. Par exemple, le CCA est beaucoup plus réceptif au feedback positif en phase de recherche, et il le sera moins en phase de répétition. Ces données renforcent l'hypothèse selon laquelle le CCA est important dans l'évaluation des actions.

Dans une tâche de Simon adaptée au singe, Nakamura *et al.* (2005) ont enregistré les neurones du CCA. Le but de leur étude était de chercher à observer la décharge des neurones du CCA dans une tâche manipulant la compatibilité. L'hypothèse testée est une des hypothèses du conflit (Botvinick *et al.*, 2001). Nakamura *et al.* (2005) n'observent pas d'augmentation de décharge du CCA dans les essais incompatibles, ce qui contredit le modèle du conflit. Nous y reviendrons dans le chapitre sur les modèles.

### 1.2.2 Ganglions de la Base

Les Ganglions de la Base (GDB) sont un ensemble de quatre noyaux sous-corticaux : le striatum, le globus pallidus, le noyau subthalamique et la substance noire (*pars compacta* et *pars reticulata*), reliés aux cortex prémoteur et préfrontal ainsi qu'au thalamus. Ils participent au contrôle des mouvements. Leur dysfonctionnement induit des troubles moteurs avec production de mouvements anormaux (hyperkinésie), de mouvements lents (akinésie, bradykinésie) et de tremblements comme dans la maladie de Parkinson. Ils ont également un rôle dans la cognition, l'humeur et les comportements non moteurs. Nous nous intéressons ici à l'implication des GDB dans les processus d'apprentissage.

Un certain nombre de travaux menés chez le singe ont montré une implication du système



mésencéphalique dopaminergique (sDA) dans les apprentissages par renforcement. Le sDA est la porte d'entrée des GDB. Il est composé de la substance noire compacte (SNc) et de l'aire tegmentale ventrale (ATV) qui se projettent sur les GDB et le cortex (Berger *et al.*, 1991; Haber et Fudge, 1997; Williams et Goldman-Rakic, 1993).

Suite à l'observation des variations de la fréquence de décharge de ces neurones au cours de l'apprentissage, Schultz et collaborateurs ont proposé que les neurones du sDA soient la clé de l'apprentissage par renforcement (Schultz, 1986; Romo et Schultz, 1990; Schultz *et al.*, 1993; Mironowicz et Schultz, 1994, 1996). Au début de l'apprentissage, l'arrivée d'une récompense non prévisible est suivie d'une augmentation phasique de la fréquence de décharge des neurones du sDA (voir le graphique du haut sur la figure 1.7). Lorsque l'animal a acquis le conditionnement, l'augmentation phasique de la fréquence de décharge des neurones ne se produit plus au moment de la récompense mais juste après l'apparition du stimulus conditionnel (voir le graphique du milieu sur la figure 1.7). Ce stimulus indique aux animaux que la récompense va être délivrée. Dans certains essais où la récompense n'a pas été administrée, les auteurs ont observé une diminution phasique de la décharge des neurones en dessous du seuil de base de décharge au moment où le singe aurait dû recevoir la récompense (voir le graphique du bas sur la figure 1.7). Ces résultats suggèrent que l'activité dopaminergique code la prédiction de la récompense. Plus exactement, elle coderait la différence entre la récompense obtenue et la récompense désirée.

Dans le champ de l'intelligence artificielle, un signal d'erreur particulier appelé "la différence temporelle de l'erreur" (TD erreur; en anglais "*temporal difference error*") est utilisé pour permettre à un algorithme d'apprendre une tâche. Cette méthode est appelée la "méthode des différences temporelles" (en anglais : "*method of temporal differences*") (Sutton, 1988; Barto, 1995; Sutton *et al.*, 1992; Sutton et Barto, 1998). À chaque essai, le réseau donne une réponse et, en parallèle, prédit la récompense qu'il va obtenir. Au moment de l'arrivée de la récompense, l'écart (ou "le signal d'erreur") entre la récompense perçue et la récompense attendue est calculé. Ce signal d'erreur peut être positif, négatif ou nul. Il est utilisé pour améliorer, à la fois, les performances du réseau et la prédiction de la récompense. Si l'apprentissage est efficace, l'écart entre la récompense espérée et la récompense reçue doit diminuer au cours des essais. Cet algorithme d'apprentissage semble fonctionner comme les neurones du sDA. Le rapprochement de ces deux observations, l'une dans le champ des neurosciences et l'autre dans le champ de l'intelligence artificielle, a permis de faire émerger l'hypothèse selon laquelle le sDA pourrait véhiculer ce signal erreur (Schultz *et al.*, 1997). Plus précisément, le sDA sensible aux récompenses informerait les GDB et le cortex frontal lorsqu'une erreur de prédiction de la récompense aurait été commise. Ces neurones émettent un

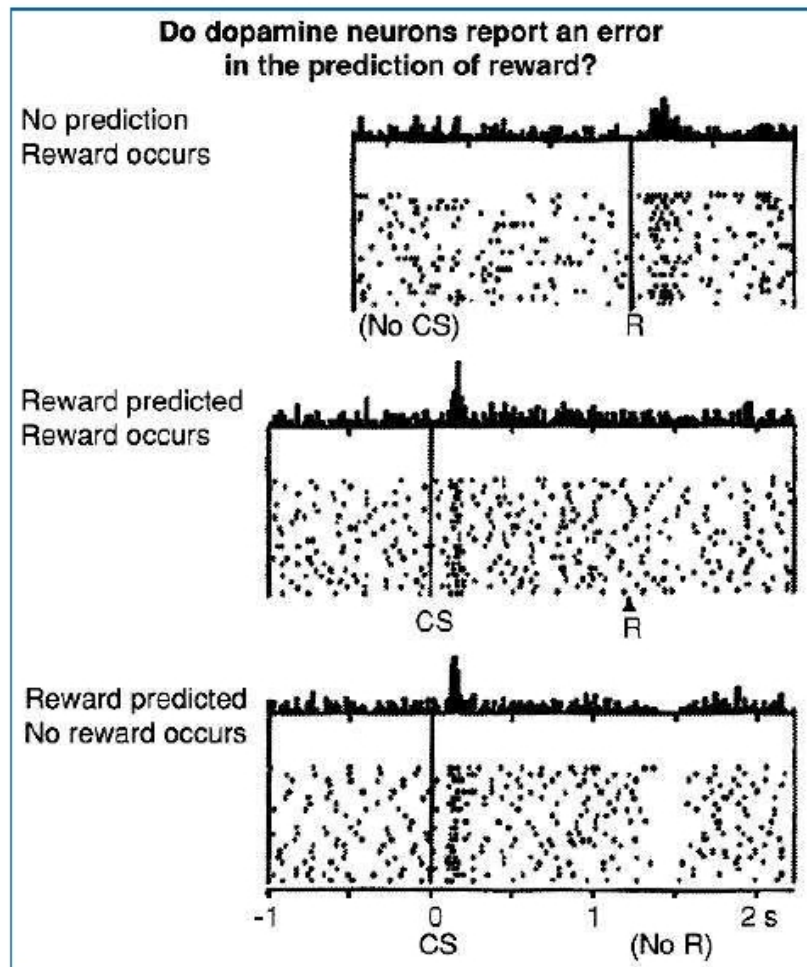


FIGURE 1.7 – Variations de la fréquence de décharge des neurones DA au moment de l'attribution de la récompense chez le primate. Figure tirée de Schultz *et al.* (1997). Ces trois graphiques montrent les variations de fréquence de décharge des neurones DA appartenant au sDA au cours d'un essai. Le graphique du haut montre l'augmentation phasique de l'activité des neurones DA avant l'apprentissage au moment où une dose de jus de fruit, non prévisible, est donnée à l'animal (R pour "reward"). Le graphique du milieu, montre une augmentation phasique de la fréquence de décharge des neurones suite à l'apparition du stimulus conditionnel (CS), une fois que l'animal a appris que ce stimulus était toujours suivi de la dose de jus de fruit (R). Le graphique du bas montre une diminution phasique de la fréquence de décharge des neurones au moment où l'animal aurait dû obtenir la dose de jus de fruit (no R).

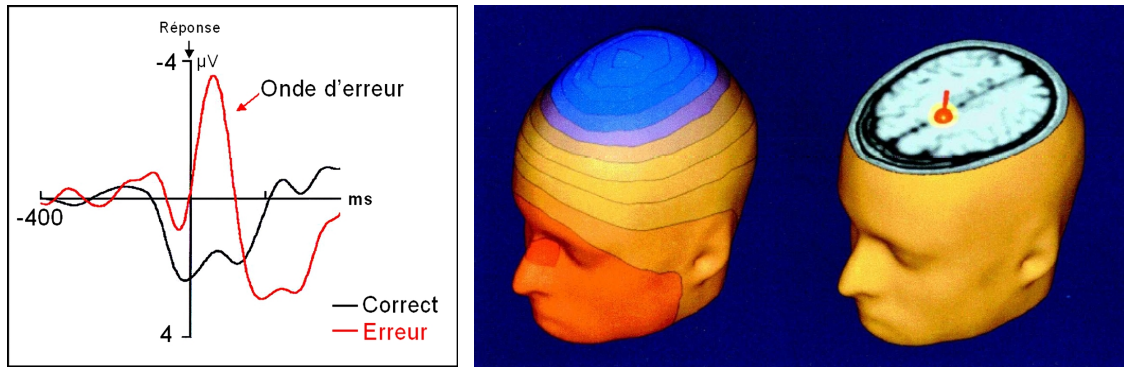


FIGURE 1.8 – **Dynamique temporelle, topographie et localisation de onde d’erreur ( $N_E$ /ERN).** Les deux illustrations sont adaptées de Van Veen et Carter (2002). Sur le graphique de gauche, il s’agit de l’activité EEG monopolaire moyennée au dessus du CCA/AMS en fonction du temps. Le zéro des temps correspond au moment où la réponse a été déclenchée. Les tracés noir et rouge correspondent respectivement à l’activité moyennée des essais corrects et des essais incorrects. Sur l’illustration de droite, nous avons une représentation de la distribution de l’activité électrique à la surface du crâne au moment du maximum d’activité de la  $N_E$  dans les erreurs (le bleu représente une activité électrique négative et le rouge une activité électrique positive). La figure de droite illustre la position du dipôle équivalent rendant compte de l’activité électrique au moment du maximum de la  $N_E$  dans les erreurs.

signal positif lorsque l’évènement est meilleur que prévu, un signal négatif lorsque l’évènement est moins bon que prévu et aucun signal si l’attente de la récompense correspond à la récompense perçue. Le rôle principal de ce système serait de faciliter le développement de programmes moteurs adaptés.

Ces aspects ont été utilisés pour la conception de plusieurs modèles neuromimétiques qui seront exposés dans le chapitre 2.3.

## 1.3 Indices EEG

### 1.3.1 L’onde d’erreur ( $N_E$ ou ERN)

Quand les sujets commettent une erreur dans une tâche de TR de choix, on observe une onde de polarité négative dans l’activité électroencéphalographique (EEG), de distribution fronto-centrale, et qui culmine environ 70 ms après la réponse mécanique du sujet ou 100 ms après le début de l’activité électromyographique (figure 1.8). Initialement observée dans les erreurs, cette activité, appelée “Onde d’erreur” ou “Négativité d’erreur” (*Error Negativity*,  $N_E$ ) par Falkenstein *et al.* (1991), a été redécouverte par Gehring *et al.* (1993) et nommée “négativité liée à l’erreur” (*Error Related Negativity*, *ERN*). Elle a été interprétée comme étant le corrélat électrophysiologique de

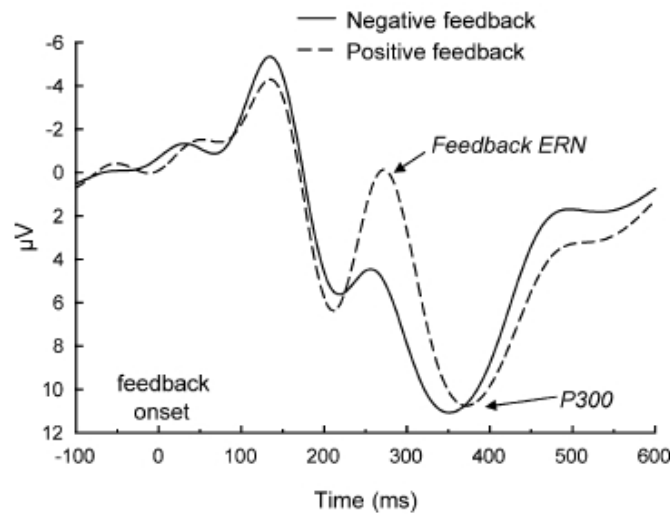


FIGURE 1.9 – **Dynamique temporelle de la négativité liée au feedback (FRN)**. Il s'agit de l'activité électrique moyenne en fonction du temps au moment où le sujet prend connaissance de sa performance à l'essai venant de se dérouler. Le zéro des temps correspond au moment d'apparition du feedback. Les tracés pleins et en pointillés correspondent respectivement à l'activité moyennée lors de l'attribution d'un feedback positif et d'un feedback négatif. Figure tirée de Nieuwenhuis *et al.* (2004).

la détection de l'erreur.

Cette activité électroencéphalographique a été localisée dans le Cortex Cingulaire Antérieur (CCA) et/ou l'Aire Motrice Supplémentaire (AMS) (Dehaene *et al.*, 1994; Van Veen et Carter, 2002; Herrmann *et al.*, 2004; Ladouceur *et al.*, 2007; O'Connell *et al.*, 2007). Un équivalent magnétique de la  $N_E$  semble avoir été observé en magnétoencéphalographie (Miltner *et al.*, 2003; Stemmer *et al.*, 2004).

### 1.3.2 La négativité liée au feedback (FRN)

Dans des situations d'apprentissage par renforcement, dans lesquelles on demande aux sujets de découvrir une règle, chaque essai est suivi d'un feedback qui informe le sujet de sa performance. Ce feedback est positif lorsque la bonne réponse a été donnée, et négatif lorsque le sujet a commis une erreur. Une activité spécifique au feedback négatif est enregistrée quelle que soit la modalité de présentation du feedback négatif. Il s'agit d'un potentiel évoqué négatif qui se développe au niveau fronto-central et qui ressemble sous plusieurs aspects à la  $N_E$  (figure 1.9). Sa latence est plus grande que celle de la  $N_E$  (le pic maximum apparaît autour de 230-330 ms après l'apparition du feedback alors que le pic maximum de la  $N_E$  apparaît 50-100 ms après la réponse mécanique

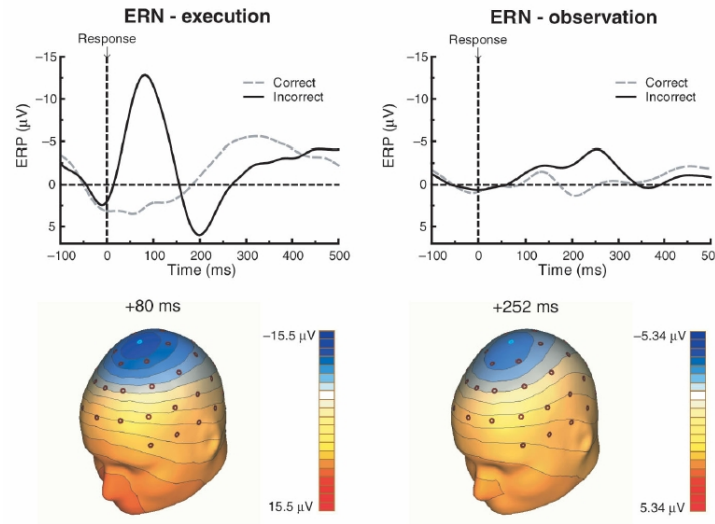


FIGURE 1.10 – **Dynamique temporelle et topographie de la “négativité de l’observateur”**. Il s’agit d’une comparaison chez les mêmes sujets de l’activité électrique cérébrale lors de la réalisation d’une tâche (graphique de gauche) et lors de l’observation d’un tiers réalisant cette même tâche (graphique de droite). Le zéro des temps correspond au moment où la réponse a été déclenchée. Les tracés en pointillés et pleins correspondent respectivement à l’activité moyennée des essais corrects et des essais incorrects. En bas, nous avons une représentation de la distribution de l’activité électrique à la surface du crâne au moment du maximum d’activité de la  $N_E$  et de la  $N_E$  de l’observateur dans les erreurs (252 et 80 ms post-réponse). Le bleu représente une activité électrique négative et le rouge une activité électrique positive. Figure tirée de Van Schie *et al.* (2004).

du sujet) et son amplitude est un peu plus petite. Cette activité a été interprétée comme étant une  $N_E$  liée au feedback négatif et appelée FRN pour “Feedback Related Negativity” (Miltner *et al.*, 1997; Ruchow *et al.*, 2002). La FRN et la  $N_E$  sont deux potentiels évoqués qui semblent refléter les mêmes processus physiologiques (Gentsch *et al.*, 2009).

Dans une étude menée en IRMf, Holroyd *et al.* (2004) ont testé l’hypothèse selon laquelle le CCA est sensible aux deux types d’informations véhiculant l’erreur : l’information de source interne (lorsque le sujet connaît la règle à suivre) et l’information de source externe (le feedback lorsque le sujet ne connaît pas la règle à suivre). Les auteurs ont observé que la partie dorsale du CCA est davantage activée dans les erreurs que dans les essais corrects pour les deux types d’erreurs.

### 1.3.3 La $N_E$ de l’observateur

Si ces activités,  $N_E$  et FRN, sont des manifestations différentes d’un même phénomène physiologique reliées à l’évaluation des performances, alors il est envisageable qu’une activité de même type soit observée chez un observateur regardant une tierce personne commettre une erreur. Milt-

ner *et al.* (2004) ont montré qu'une  $N_E$  se développait chez les sujets qui observaient une erreur commise par une simulation d'une tâche de temps de réaction par un ordinateur. Des résultats similaires ont été retrouvés chez le sujet observant un humain effectuer la tâche (Van Schie *et al.*, 2004; Bates *et al.*, 2005). La  $N_E$  et la  $N_E$  de l'observateur ont été comparée par Van Schie *et al.* (2004) (voir figure 1.10). La topographie des deux activités semble similaire mais la dynamique temporelle de la  $N_E$  de l'observateur semble plus proche de la dynamique temporelle de la FRN que de celle de la  $N_E$ .

### 1.3.4 La $N_E$ , la FRN et la $N_E$ de l'observateur

Ces trois activités pourraient avoir comme dénominateur commun la fonction d'évaluation.

- La  $N_E$  apparaît suite à une évaluation interne de sa propre erreur.
- La FRN apparaît suite à une évaluation externe de son erreur.
- La  $N_E$  de l'observateur apparaît à la suite de l'évaluation interne d'une erreur faite par autrui.

Toutefois, des différences et des similitudes apparaissent entre la  $N_E$  de l'observateur, la  $N_E$  et la FRN. La latence de la  $N_E$  de l'observateur est beaucoup plus longue que celle de la  $N_E$  (252 versus 80 ms dans l'étude de Van Schie *et al.*, 2004). Notons que la latence de la  $N_E$  de l'observateur est comparable à celle de la FRN se situant entre 230 et 330 ms. D'autre part, l'amplitude de la  $N_E$  de l'observateur est plus petite que celle de la  $N_E$ . Cette observation est également valable pour la comparaison entre la FRN et la  $N_E$ . Gentsch *et al.* (2009) ont apporté un argument supplémentaire en montrant que  $N_E$  et la FRN sont générés par une même source. Il semblerait qu'il existe un mécanisme général lié à la gestion de l'erreur localisé au niveau du CCA que cette erreur soit effectuée, dénoncée ou observée.

### 1.3.5 La Positivité d'Erreur ( $P_E$ )

La "Positivité d'Erreur" (ou "Error Positivity") est un potentiel évoqué par la réponse incorrecte (Falkenstein *et al.*, 1991) que l'on observe après la  $N_E$  et qui culmine environ 300 ms après le début de l'activité électromyographique (voir figure 1.11). Sa topographie est davantage postérieure que celle de la  $N_E$ .

La  $P_E$  est spécifique des erreurs puisqu'elle n'apparaît ni dans le cas des réponses correctes, ni dans le cas des ébauches d'erreur (Vidal *et al.*, 2000). Elle n'est sensible ni à la compatibilité, ni à la consigne vitesse/précision (Bartholow *et al.*, 2005). Nieuwenhuis *et al.* (2001) ont montré

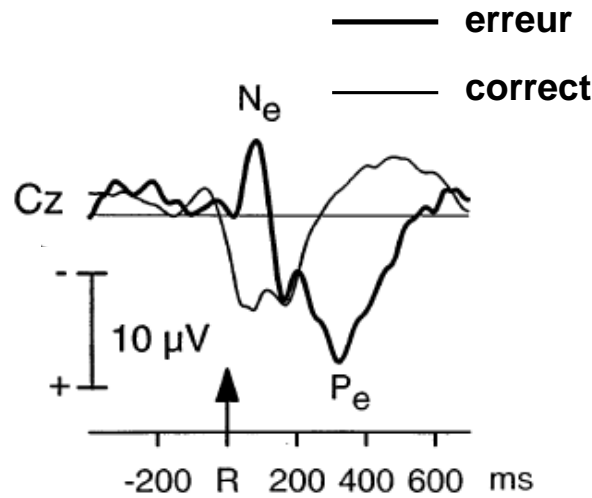


FIGURE 1.11 – **Dynamique temporelle de la Positivité d'Erreur ( $P_E$ )**. Il s'agit de l'activité EEG monopolaire moyennée pour l'électrode Cz en fonction du temps. Le zéro des temps correspond au moment où la réponse a été déclenchée. Les tracés fin et épais correspondent respectivement à l'activité moyennée des essais corrects et des essais incorrects. La  $P_E$ , uniquement présente dans les erreurs, se développe après la  $N_E$  et pique entre 300 et 400 ms après la réponse. Figure tirée de Falkenstein *et al.* (1991).

que, contrairement à la  $N_E$ , la  $P_E$  n'est présente que pour les erreurs conscientes et que seules les erreurs conscientes sont suivies d'un ralentissement post-erreur. Ces résultats ont été confirmés par d'autres études (Hajcak *et al.*, 2003; O'Connell *et al.*, 2007; Endrass *et al.*, 2005, 2007). La  $P_E$  semble donc être le corrélât électrophysiologique de la prise de conscience de l'erreur. Cette théorie a toutefois été remise en cause par Overbeek *et al.* (2005) qui défendent l'idée que la  $P_E$  est une composante émotionnelle.

### 1.3.6 Activation et Inhibition des cortex moteurs

Taniguchi *et al.* (2001) ont étudié la dynamique temporelle des cortex moteurs dans la période préparatoire et pendant l'exécution de la réponse manuelle dans une tâche de TR simple et une tâche de TR double choix. Les auteurs ont mis en évidence l'existence de deux ondes qui se développent au dessus des cortex moteurs primaires juste avant le début de la réponse motrice des sujets. Cette observation a été rendue possible grâce au calcul du Laplacien de surface et confirmée par d'autres études (Vidal *et al.*, 2003b; Tandonnet *et al.*, 2003).

Dans une tâche de TR de choix bimanuel, on observe, en regard de l'aire motrice primaire (M1) controlatérale à la réponse, une onde de polarité négative qui se développe environ 100 ms

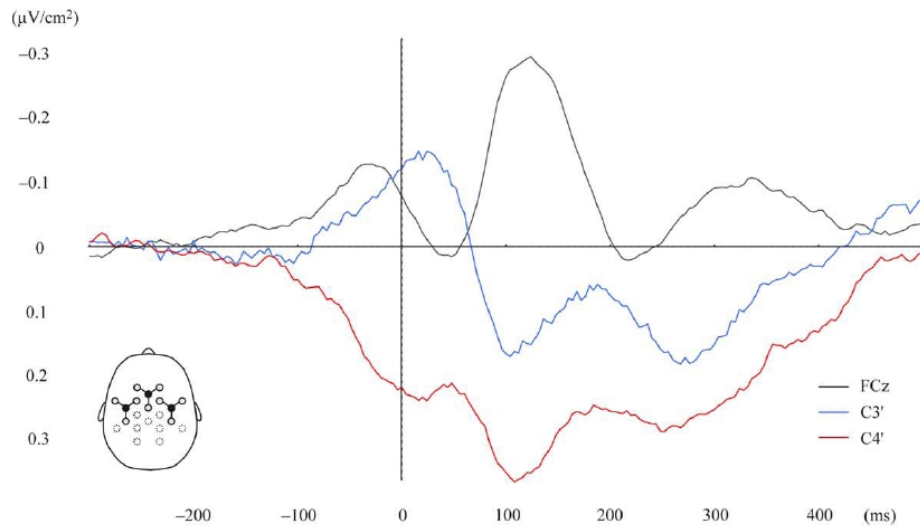


FIGURE 1.12 – **Dynamique temporelle de l'activité EEG après calcul du Laplacien de surface au dessus des cortex moteurs (tracés rouge et bleu) et au dessus de l'AMS/CCA (tracé noir).** Les tracés bleu et rouge correspondent respectivement à l'activité du cortex moteur impliqué dans la réponse (cortex controlatéral - électrode C3) et du cortex moteur non impliqué dans la réponse (cortex ipsilatéral - électrode C4). La N-40 est visible sur le tracé noir avant le début de l'EMG (électrode FCz). Figure tirée de Tandonnet *et al.* (2005).

avant le début de l'EMG et qui culmine autour du début de l'EMG (voir le tracé bleu sur la figure 1.12). Dans le même temps, une onde positive se développe en regard de l'aire motrice primaire ipsilatérale à la réponse (voir le tracé rouge sur la figure 1.12). Les auteurs ont interprété l'onde négative (observée sur M1 controlatérale à la réponse) comme étant le reflet de l'activation du cortex moteur impliqué dans la réponse, et l'onde négative (observée sur M1 ipsilatérale à la réponse) comme étant le reflet de l'inhibition du cortex moteur impliqué dans la réponse alternative.

Cette interprétation va dans le sens des résultats observés dans deux autres études s'intéressant à l'organisation de la commande motrice au niveau spinal (en étudiant le réflexe H) et cortico-spinal (en étudiant l'excitabilité de la voie cortico-spinale par stimulation magnétique transcrânienne) (Hasbroucq *et al.*, 2000; Burle *et al.*, 2002a). Ces différentes techniques ont montré que juste avant la réponse, l'excitabilité de la région cortico-spinale impliquée dans cette commande volontaire augmentait, tandis que, au même moment celle de la région cortico-spinale non-impliquée diminuait (pour une revue voir Burle *et al.*, 2004b).

Cette inhibition des structures cortico-spinales impliquées dans la commande volontaire reflète très probablement un système de prévention de l'erreur. Pour éviter de se tromper dans les tâches de TR de choix bimanuel, la commande motrice serait caractérisée par deux actions : d'une part



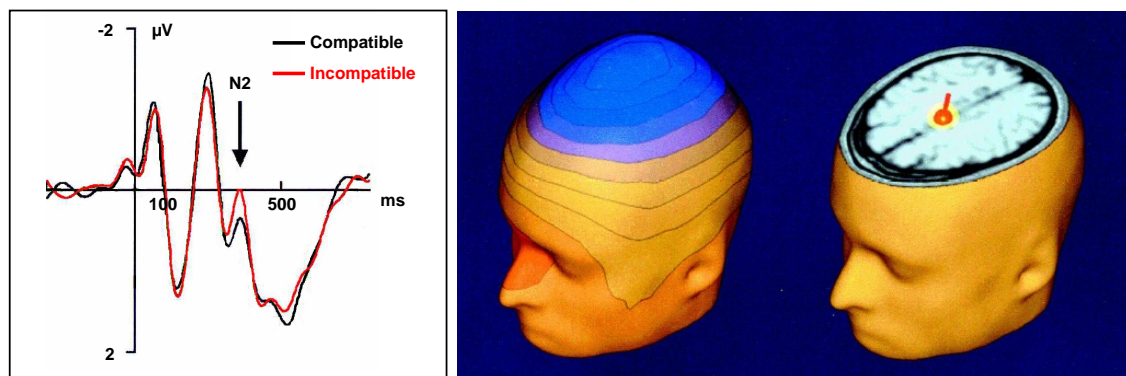


FIGURE 1.13 – **Dynamique temporelle de la  $N2$  calée sur le stimulus.** Les tracés noir et rouge correspondent respectivement à l'activité dans les essais compatibles et incompatibles corrects. Le zéro des temps correspond au moment de présentation du stimulus. Sur la figure de droite est représenté la localisation dipolaire de l'activité  $N2$ . Figures adaptées de Van Veen et Carter (2002).

l'activation de la structure motrice impliquée dans la réponse et d'autre part l'inhibition de la structure motrice impliquée dans la réponse alternative. Cette inhibition permettrait de diminuer le risque d'erreur et pourrait être le résultat de la mise en place d'un contrôle de l'action. D'ailleurs Carbonnell *et al.* (2004) ont montré que cette inhibition est réduite voire absente dans les tâches de TR simple, ce qui suggère qu'elle est sensible au contexte alors que l'activation ne l'est pas.

### 1.3.7 La N-40

En étudiant la dynamique des activations des cortex moteurs pendant la période préparatoire et l'exécution de la réponse, Vidal *et al.* (2003b) ont observé une onde de polarité négative qui se développe en regard de l'AMS (au dessus de l'électrode FCz). Elle débute environ 200 ms avant le début de l'activité EMG et atteint son maximum environ 40 ms avant le début de l'EMG (voir le tracé noir de la figure 1.12). Cette onde a été nommée "N-40". (Vidal *et al.*, 2003b) ont montré que cette onde se développait avant le patron d'activation-inhibition observé au dessus des cortex moteurs, ce qui suggère un lien hiérarchique entre la N-40 et le patron d'activation-inhibition. Son rôle n'est pas encore bien établi.

### 1.3.8 La $N2$

Les partisans du modèle du conflit que nous expliquerons dans le chapitre suivant, défendent que la  $N_E$  observée dans les erreurs reflète un conflit de réponses. Ils proposent également que dans les essais corrects, le conflit de réponse, plus important dans les essais incompatibles que les essais

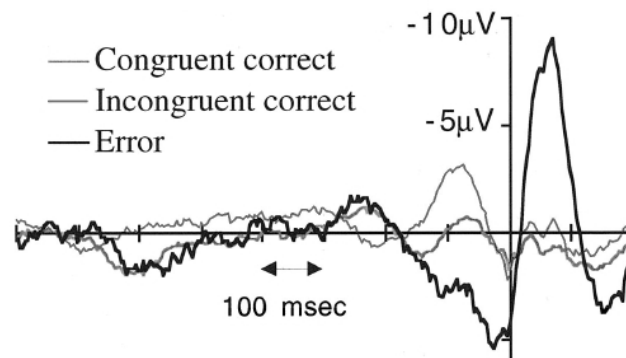


FIGURE 1.14 – **Dynamique temporelle de la  $N_2$  calée sur la réponse.** Il s'agit des tracés moyennés au niveau de l'électrode FCz des différents types d'essais calés sur la réponse mécanique du sujet. Les tracés gris clair, gris foncé et noirs correspondent à l'activité EEG respectivement dans les essais compatibles corrects, incompatibles corrects et dans les erreurs. Figure adaptée de Yeung *et al.* (2004).

compatibles, se produit avant la réponse (Van Veen et Carter, 2002; Ladouceur *et al.*, 2007).

Or dans les essais corrects, ils observent une onde de polarité négative au niveau fronto-central qui est plus ample dans les essais incompatibles que dans les essais compatibles (figure 1.13). Cette activité est nommée  $N_2$ . Yeung *et al.* (2004) ont observé une modulation d'une onde de polarité négative avant la réponse dans les tracés calés sur la réponse correcte. Cette  $N_2$  a été interprétée comme reflétant le conflit dans les essais corrects (figure 1.14).

La  $N_2$  calée sur la réponse et la  $N_{-40}$  se ressemblent beaucoup. Nous remarquons que la  $N_{-40}$  a été observée dans les essais dits "purs-corrects" dans lesquels aucune activation EMG n'a été observée précédant la bouffée EMG ayant donné la réponse. Or, comme nous le verrons dans le paragraphe suivant, on observe une  $N_E$  au moment des ébauches d'erreur (voir figure 1.16). Puisque les ébauches d'erreur sont plus fréquentes dans les essais incompatibles que dans les essais compatibles (Eriksen *et al.*, 1985), la  $N_2$ , calée sur le stimulus, observée dans tous les essais corrects pourrait être le reflet de la  $N_E$  déclenchée au moment de l'ébauche d'erreur. De la même façon, la  $N_2$  observée calée sur la réponse, pourrait être due à la  $N_E$  déclenchée par l'ébauche d'erreur.

## 1.4 La spécificité de l'Onde d'Erreur

Comme nous l'avons déjà précisé ci-dessus, l'onde d'erreur a été initialement observée juste après la production d'une erreur et sa source a été localisée dans le CCA et/ou l'AMS. Depuis la découverte de son existence, cette onde a suscité l'intérêt de nombreuses équipes de recherche. On

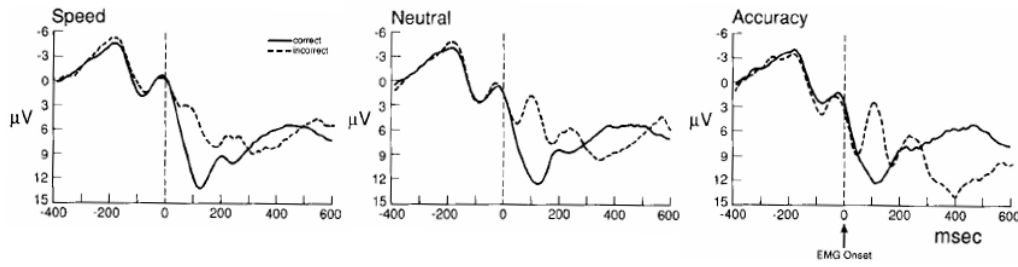


FIGURE 1.15 – **Grand moyennage des tracés EEG monopolaires pour les erreurs (traits pointillés) et les essais corrects (traits pleins) en fonction de la consigne vitesse ou précision.** Les sujets doivent privilégier soit la vitesse (graphique de droite), soit la précision (graphique de gauche) ou bien doivent maintenir un équilibre entre les deux (graphique central). La taille de la  $N_E$  dans les erreurs varie en fonction de la consigne : l’amplitude de la  $N_E$  des erreurs augmente avec la précision accordée à la tâche. Figure tirée de Gehring *et al.* (1993).

a découvert qu’elle était très certainement le résultat d’un processus global puisqu’elle ne semble ni spécifique à une modalité sensorielle de présentation des stimuli, ni spécifique à une modalité de réponse. Falkenstein *et al.* (1991) ont montré qu’une  $N_E$  était observée dans les erreurs suite à une présentation des stimuli dans la modalité visuelle et la modalité auditive. On l’observe également quelle que soit la modalité de réponse, dans les tâches où les réponses devaient être produites avec les mains (Falkenstein *et al.*, 1991; Gehring *et al.*, 1993), les pieds (Holroyd *et al.*, 1998; Gehring et Fencsik, 2001), impliquant l’articulation vocale de mots (Masaki *et al.*, 2001), et l’exécution de saccades oculaires (Van ’t Ent et Apkarian, 1999; Nieuwenhuis *et al.*, 2001).

L’amplitude de la  $N_E$  semble révélatrice de l’importance que le sujet accorde à l’erreur. Plus on demande aux sujets d’être précis, plus les TR sont longs et plus le nombre d’erreurs est faible. Inversement, plus on demande aux sujets d’être rapides, plus les TR sont courts et plus le nombre d’erreurs est important. Cet effet est connu sous le nom d’échange vitesse-précision (“speed-accuracy tradeoff”). Gehring *et al.* (1993) ont montré que l’amplitude de la  $N_E$  dans les erreurs suit la fonction d’échange vitesse-précision : plus on demande aux sujets d’être précis, plus l’amplitude de la  $N_E$  est grande (figure 1.15).

#### 1.4.1 La $N_E$ est-elle spécifique aux erreurs ?

Initialement, la  $N_E$  a été observée uniquement lors de la production d’erreurs (Falkenstein *et al.*, 1991; Gehring *et al.*, 1993). Dans certains essais corrects, il arrive parfois qu’une petite activité EMG incorrecte, appelée “ébauche d’erreur”, ait été produite avant l’activité EMG qui a déclenché la réponse correcte (figure 1.5 page 16). Après moyennage des tracés EEG calés sur le

début de l'ébauche d'erreur, on observe une  $N_E$  dont l'amplitude est plus petite que celle observée dans les erreurs (Scheffers *et al.*, 1996; Vidal *et al.*, 2000).

Vidal *et al.* (2000) ont montré qu'il existait une négativité dans les essais corrects dont la polarité, la morphologie, la latence et la topographie étaient similaires à la  $N_E$  observée dans les erreurs (voir figure 1.16). Sa présence a suscité un débat important puisque dans de nombreuses études elle n'était pas observée. Son existence a explicitement été remise en cause par Coles *et al.* (2001). Vidal et collaborateurs expliquent qu'ils ont pu observer la négativité dans les essais corrects grâce au calcul du Laplacien de surface. Cette technique améliore la résolution spatiale de l'EEG<sup>2</sup>. Répondant à la principale critique de Coles *et al.* (2001), les auteurs ont montré que la négativité dans les essais corrects n'était pas le fruit d'un artefact de moyennage (Vidal *et al.*, 2003a).

Aujourd'hui l'existence de la négativité dans les essais corrects est admise par le plus grand nombre, mais la question de son origine se pose toujours. Cette négativité observée dans les essais corrects a-t-elle la même origine fonctionnelle que la  $N_E$  observée dans les erreurs ? Cette question précise a fait d'objet d'une étude à part entière dans cette thèse dans le chapitre 4.

#### 1.4.2 La taille de la $N_E$ varie-t-elle en fonction du nombre d'éventualités ?

Les premières études se sont focalisées sur la  $N_E$  des erreurs. Bernstein *et al.* (1995) et Falkenstein *et al.* (2000) ont utilisé une tâche à 4 éventualités impliquant 4 effecteurs : 2 doigts sur chacune des mains. Les résultats montrent que la  $N_E$  des erreurs est plus grande lorsque l'erreur est double (erreur de main et erreur de doigt) en comparaison avec une erreur simple (erreur de main ou de doigt). Ils observent aussi une différence entre les erreurs simples : la  $N_E$  des erreurs est plus petite dans les erreurs de doigt que dans des erreurs de main. Ces résultats ont été interprétés selon l'hypothèse de la mismatch (détaillée au paragraphe 2.2.1 du chapitre 2) qui stipule que la  $N_E$  reflète la différence entre la réponse exécutée et la représentation de la réponse correcte.

Ces résultats ont été remis en cause par Gehring et Fencsik (2001) qu'ils n'estimaient pas interprétables puisque les réponses "doigt" d'une même main (index et majeur) ont une partie de leurs programmes moteurs en commun. En effet, les doigts sont trop proches et nécessitent

---

2. Si l'on admet que le milieu est homogène et isotrope, appliquer l'opérateur Laplacien consiste à calculer la somme des doubles dérivées partielles :  $\Delta V(x, y) = \frac{\partial^2 V(x, y)}{\partial x^2} + \frac{\partial^2 V(x, y)}{\partial y^2}$ , où  $\Delta V$  est le Laplacien du potentiel électrique (exprimé classiquement en EEG en  $\mu V/cm^2$ ),  $x$  et  $y$  sont les coordonnées cartésiennes du plan tangentiel à la surface du cuir chevelu, et  $V$  est une fonction de 2 variables  $x$  et  $y$  décrivant la distribution spatiale du potentiel. Le Laplacien de surface est interprétable physiquement : il reflète la composante radiale de la densité de courant. Il est également interprétable physiologiquement puisqu'il approxime le corticogramme (activité électrique en surface du cortex).

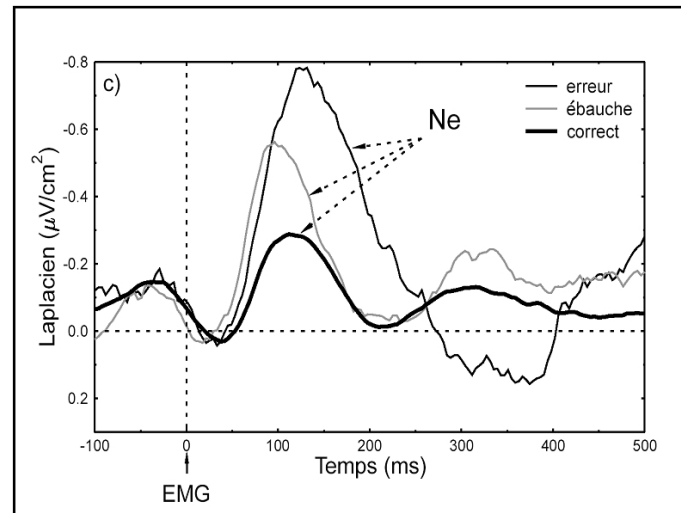


FIGURE 1.16 – Grands moyennages des tracés EEG après calcul du Laplacien de surface, calés sur le début de l’EMG pour les essais corrects (ligne épaisse), les essais corrects avec ébauche d’erreur (ligne fine) et les erreurs (ligne intermédiaire). Le zéro des temps correspond au début de l’activité musculaire de la bouffée EMG correcte dans les essais corrects et de la bouffée EMG incorrecte pour les erreurs et les ébauches d’erreur. La taille de l’onde d’erreur varie selon la nature des essais. Dans ces essais corrects, l’onde d’erreur a la plus faible amplitude, et cette amplitude augmente dans les essais corrects avec ébauche d’erreur, puis dans les erreurs. Adapté d’après Vidal *et al.* (2000).

l’utilisation de mêmes muscles pour mettre en œuvre le mouvement. En utilisant une tâche de Stroop impliquant 4 effecteurs différents (main ou pied ; droit ou gauche), Gehring et Fencsik (2001) ont distingué les erreurs “ipsilatérales” (bon côté mais erreur sur l’effecteur), “controlatérales” (bon effecteur mais erreur sur le côté de réponse) et les erreurs “opposées” (erreur à la fois sur le côté de réponse et sur l’effecteur). Au niveau comportemental, les sujets font davantage d’erreurs “ipsilatérales” que “controlatérales” ; les erreurs “opposées” sont les plus rares. L’amplitude de la  $N_E$  est la plus grande dans les erreurs “ipsilatérales” puis dans les erreurs “controlatérales” et enfin dans les erreurs “opposées”. Ces résultats ont été interprétés selon l’hypothèse du conflit (voir paragraphe 2.2.3). Le conflit de réponse serait plus important dans les erreurs “ipsilatérales” que “controlatérales” ainsi que dans les erreurs dites “opposées”.

Vidal *et al.* (2000) ont montré que la négativité observée dans les essais corrects dans une tâche de TR à deux éventualités est également présente dans une tâche de TR simple. Cela suggère qu’elle n’est pas liée à l’incertitude de la réponse ni à la précision de la réponse. À notre connaissance, peu d’études se sont directement intéressées aux variations d’amplitude de la négativité dans les essais corrects en fonction du nombre d’éventualités de réponses. Dans le but d’étudier le développement

des processus de contrôle, Hogan *et al.* (2005) se sont intéressés à l'effet de la complexité de la tâche sur la négativité dans les essais corrects et les erreurs en fonction de l'âge (adolescents/adultes). Les tâches dites "simple" et "complexe" étaient respectivement des tâches de TR à deux et à quatre éventualités. Dans la tâche simple, une flèche qui pointe vers la gauche ou vers la droite. Le sujet doit appuyer sur le bouton de réponse situé du côté vers lequel pointe la flèche. Dans la tâche complexe, la flèche apparaît en vert ou en rouge et indique au sujet s'il doit répondre de façon compatible ou incompatible à la direction de la flèche. La compatibilité stimulus-réponse est manipulée uniquement dans la tâche à 4 éventualités. Dans la tâche simple, la  $N_E$  des erreurs est plus ample que la négativité dans les essais corrects chez les adolescents et les adultes. En revanche, cette différence n'existe plus dans la tâche complexe chez les adolescents mais persiste chez les adultes. Chez les adolescents, il y a une augmentation d'amplitude de la négativité des corrects et une diminution de la  $N_E$  des erreurs dans la tâche complexe par rapport à la tâche simple que le stimulus soit compatible ou incompatible. Chez les adultes, la  $N_E$  des erreurs est beaucoup plus ample dans la condition incompatible que dans la condition compatible. Il est assez difficile de conclure sur cette étude pour plusieurs raisons : (1) il ne s'agit pas réellement de la manipulation du nombre d'éventualités puisque dans les deux tâches, deux réponses sont proposées. Il s'agit plus d'une complexité au niveau des règles d'association S-R. (2) les essais compatibles et incompatibles ne sont pas équiprobables (25% pour les incompatibles et 75% pour les compatibles). Il y a donc potentiellement un facteur confondu avec la rareté des stimuli incompatibles. Ces effets sont-ils dus à l'augmentation de la difficulté de la tâche, ou bien sont-ils la conséquence de la manipulation de la probabilité d'occurrence des essais compatibles et incompatibles ? (3) il s'agit avant tout d'une étude développementale, l'interaction entre l'âge et la complexité de la tâche est très intéressante. Elle suggère que le CCA est une structure en cours de développement chez l'enfant et qui met en lumière la maturation des systèmes de contrôle.

Au vu de ces études, il est difficile de connaître l'effet du nombre d'éventualités sur l'amplitude de la  $N_E$  dans les erreurs et les essais corrects. Les résultats sur la  $N_E$  des erreurs sont contradictoires (Bernstein *et al.*, 1995; Falkenstein *et al.*, 2000; Gehring et Fencsik, 2001) et ceux sur l'amplitude de la négativité dans les essais corrects ne sont pas facilement interprétables.

### 1.4.3 La $N_E$ est-elle sensible à la compatibilité ?

Scheffers et Coles (2000) ont montré que la  $N_E$  des erreurs est plus grande dans la condition compatible que dans la condition incompatible. Dans des données publiées dans la thèse de Sonia

Allain, cet effet de compatibilité sur la  $N_E$  a été retrouvé dans les erreurs, et a été observé dans les essais corrects dans la tâche de Simon (Allain, 2005). Curieusement, les résultats issus de cette expérience ne montrent pas de différence d'amplitude de la  $N_E$  des ébauches d'erreur en fonction de la compatibilité.

#### 1.4.4 La taille de la $N_E$ peut-elle présager de la performance à l'essai suivant ?

Ridderinkhof *et al.* (2003) sont les premiers à s'être intéressés à cette question. Ils ont analysé la positivité que l'on observe dans les essais corrects au moment de la réponse correcte. Cette positivité n'est observée que dans les enregistrements EEG monopolaires (1.9). Ils ont montré que cette positivité était augmentée dans les essais corrects précédant une erreur. L'augmentation d'amplitude annonçant l'erreur à l'essai suivant (nommée "*Error Preceding Positivity*", EPP) a été interprétée comme étant le reflet d'une défaillance momentanée des processus de contrôle qui est tenue responsable de l'erreur survenant à l'essai suivant.

Puisque le calcul du Laplacien de surface sur les tracés EEG monopolaires a permis de révéler l'existence d'une négativité dans les essais corrects (Vidal *et al.*, 2000, 2003a), Allain *et al.* (2004b) se sont demandés si l'augmentation de la positivité dans les essais corrects précédents une erreur observée sur les tracés monopolaires par Ridderinkhof *et al.* (2003) n'était pas due à une diminution de l'amplitude de la négativité dans les essais corrects observable après calcul du Laplacien de surface. En accord avec cette prédiction, les auteurs montrent que l'amplitude de la négativité dans les essais corrects est diminuée dans les essais précédents une erreurs (Allain *et al.*, 2004b). Ces résultats suggèrent que la négativité dans les essais corrects est également liée aux processus de supervision de la performance au même titre que les  $N_E$  dans les erreurs et les ébauches d'erreur. Ces observations ont été confirmées récemment par Cavanagh *et al.* (2009).

Observe t-on cet effet sur la  $N_E$  des erreurs ? Cette prédiction a été testée directement par Hajcak et Simons (2008). Ils se sont intéressés à l'amplitude de la  $N_E$  des erreurs en fonction de la performance à l'essai suivant. Quelle que soit la séquence (correct-erreur-erreur ou correct-erreur-correct), l'amplitude de la  $N_E$  des erreurs au deuxième essai est la même. Ce qui suggère que l'amplitude de la  $N_E$  des erreurs ne peut pas être un bon indicateur de la performance à l'essai suivant. En revanche, l'analyse des TR à l'essai suivant l'erreur montrent que le ralentissement post-erreur est important lorsque l'essai est correct (séquence correct-erreur-correct), et il est faible lorsque le sujet produit une nouvelle erreur (séquence correct-erreur- erreur). La  $N_E$  semble signaler

la nécessité d'augmenter le contrôle cognitif mais ne semble pas être directement impliquée dans la mise en place des ajustements comportementaux. Hajcak et Simons (2008) concluent que la  $N_E$  est un substrat indirect des ajustements à la performance et que l'activité du CCA est nécessaire mais pas suffisante à la mise en place des ajustements comportementaux. Cette hypothèse est soutenue par les résultats comportementaux de patients souffrants d'une lésion du CCA (di Pellegrino *et al.*, 2007). L'effet de compatibilité est présent, mais il n'est pas modulé en fonction de la nature de l'essai précédent. Ce résultat suggère que le CCA est nécessaire à la mise en place de l'ajustement comportemental.

Les résultats de Gehring et Fencsik (2001) vont aussi dans ce sens. Les auteurs ont étudié les modulations d'amplitude de la  $N_E$  dans les erreurs en fonction de la présence ou de l'absence d'un ralentissement post-erreur enregistré dans les essais corrects qui suivent l'erreur. Aucune différence d'amplitude de la  $N_E$  n'a été notée. En revanche, les résultats observés par Ladouceur *et al.* (2007) sont opposés : après avoir trié les erreurs en fonction de faible ou fort ralentissement post-erreur qu'ils présentaient à l'essai suivant, ils observent une augmentation de l'amplitude de la  $N_E$  dans les erreurs suivies d'un fort ralentissement, et ce quel que soit l'âge des sujets (jeunes adolescents, adolescents et adultes). Ces résultats sont cohérents avec ceux de Kerns *et al.* (2004) obtenus en IRMf qui montrent une augmentation de l'activité du CCA dans les essais suivant une erreur.

La  $N_E$  des corrects dans une certaine mesure "présage" la nature (correcte ou incorrecte) de la réponse qui va suivre. Cette observation ne semble pas systématique pour le  $N_E$  des erreurs.

### 1.4.5 Conclusion

Une négativité est observée dans les erreurs, les ébauches d'erreurs et dans les essais corrects. Aujourd'hui la présence de la négativité dans les essais corrects n'est plus remise en cause (Vidal *et al.*, 2003a). Pour autant, il n'est pas admis que les trois négativités reflètent un même processus physiologique (Yordanova *et al.*, 2004). Dans le chapitre expérimental 4, nous avons cherché à apporter un argument supplémentaire en faveur de l'unicité de ces trois ondes.

## 1.5 Pathologies et Contrôle

Il est intéressant d'étudier les marqueurs comportementaux et électrophysiologiques des processus de contrôle lorsque des dysfonctionnements de ces processus ont été observés sur la base du comportement. Tout d'abord, un certain nombre de publications montrent que l'efficacité des processus de contrôle n'est pas constante tout au long de la vie. Comment se mettent en place ces



processus chez l'enfant ? Quel est l'effet du vieillissement normal sur les processus de contrôle ? Dans un deuxième temps, nous nous sommes intéressés aux différentes pathologies conduisant à des dysfonctionnements des processus de contrôle. Enfin, nous nous sommes intéressés à quelques situations expérimentales conduisant à altérer ces processus de contrôle de l'action.

### 1.5.1 Effets dûs au développement normal

Un certain nombre d'études ont comparé les marqueurs des processus de contrôle chez les enfants, les adolescents et les jeunes adultes (Falkenstein *et al.*, 2001; Davies *et al.*, 2004b,a; Santesso *et al.*, 2006; Ladouceur *et al.*, 2007). Les résultats sont très cohérents. Chez les enfants, les TR sont allongés et le nombre d'erreurs est plus grand. C'est moins le cas chez les adolescents. En revanche, le ralentissement post-erreur est toujours présent quel que soit l'âge. Au niveau électrophysiologique, la  $N_E$  des erreurs chez l'enfant est plus petite que chez l'adolescent et chez l'adulte, alors que la  $N_E$  des essais corrects est plus grande chez l'enfant que chez l'adolescent et l'adulte (Santesso *et al.*, 2006; Ladouceur *et al.*, 2007). Il n'y a pas d'effet de l'âge sur la  $P_E$ . Il semblerait que les enfants ne soient pas capables d'évaluer leurs comportements aussi bien que le font les adultes. En revanche, ils ont parfaitement conscience de leurs erreurs et parviennent à mettre en place des ajustements post-erreurs. Ce qui ressort de ces résultats est une dissociation claire entre l'amplitude de la négativité dans les erreurs et les essais corrects, et la mise en place des ajustements comportementaux. Deux hypothèses sont proposées pour expliquer ces résultats. Il pourrait s'agir d'une maturation lente du CCA ou alors d'une implication tardive du CCA dans les processus de contrôle. Cette hypothèse est corroborée par des résultats obtenus en IRMf par Adelman *et al.* (2002) qui montrent une augmentation du niveau d'activité du CCA entre l'adolescence et la période jeune adulte. La deuxième hypothèse est émotionnelle : la  $N_E$  serait reliée à la saillance de l'erreur. Plus l'erreur est importante pour le sujet, plus la  $N_E$  sera grande (Hajcak *et al.*, 2005a). Comme l'ont évoqué Segalowitz et Davies (2004), les enfants continuellement en situation d'apprentissage font beaucoup d'erreurs, ce qui expliquerait que la composante émotionnelle de l'erreur soit plus faible chez eux que chez les adultes. Néanmoins, cette interprétation n'explique pas pourquoi la négativité dans les essais corrects est augmentée.

Mathalon *et al.* (2003) ont testé les effets du vieillissement normal. Les sujets âgés sont plus lents que les sujets jeunes. Néanmoins les pourcentages de réussite sont équivalents dans les deux groupes. Au niveau électrophysiologique, les sujets âgés montrent à la fois, une réduction de la  $N_e$  des erreurs et de celle des essais corrects. La diminution des négativités observées dans les erreurs

et les essais corrects pourraient être dues au déclin du nombre de récepteurs dopaminergiques dû à l'âge (Nieuwenhuis *et al.*, 2002).

### 1.5.2 Hypofonctionnement des processus de contrôle

Vlamings *et al.* (2008) ont observé un déficit des processus de contrôle chez les enfants autistes. Les TR moyens et le taux d'erreur de ces enfants sont comparables à ceux des enfants contrôles. En revanche, ils ne présentent pas de ralentissement post-erreurs. Au niveau électrophysiologique, la  $N_E$  des erreurs est anormalement diminuée et se retrouve au même niveau que celle dans les essais corrects. Les auteurs expliquent ce résultat par une immaturité du CCA qui refléterait l'insensibilité à détecter des situations dans lesquelles le risque d'erreur est élevé. Cette hypothèse prend appui sur les données d'Haznedar *et al.* (1997) qui, chez l'autiste adulte, ont montré un métabolisme du CCA plus bas que celui chez les sujets contrôles. Vlamings *et al.* (2008) ont également noté une diminution de la  $P_E$  chez les enfants autistes. Ce résultat peut être interprété par une incapacité à prendre conscience de l'erreur ou par une réduction de l'allocation des ressources attentionnelles lorsqu'ils font des erreurs.

Dans l'étude de Mathalon *et al.* (2003), déjà évoqué au paragraphe 1.5.1, les auteurs ont comparé des sujets jeunes et sains aux sujets âgés sains, ainsi qu'avec des patients âgés atteints de la maladie d'Alzheimer. La comparaison des performances entre sujets sains âgés et patients montre des différences importantes au niveau des TR et des pourcentages de réussite. Les TR sont allongés chez les patients et le pourcentage de réussite est plus faible. Au niveau électrophysiologique, la  $N_E$  des essais corrects est diminuée et celle des erreurs est très largement diminuée. Les deux négativités ont la même amplitude. Les auteurs supposent que ces résultats sont dus à un déficit du CCA et du DLPFC, et/ou des connexions entre les deux structures.

Chez les patients schizophrènes, la négativité des erreurs est considérablement réduite tandis que celle des corrects est augmentée (Mathalon *et al.*, 2002; Bates *et al.*, 2002). Toutefois, l'administration d'un traitement médicamenteux permet de limiter la diminution de la  $N_E$  des erreurs (Bates *et al.*, 2004). En revanche, la  $P_E$  est normale. Plusieurs hypothèses explicatives sont envisagées. Des dysfonctionnements du CCA et du DLPFC ont également été proposés pour expliquer ces résultats. D'ailleurs, ces résultats peuvent être rapprochés de ceux obtenus par Gehring et Knight (2000) qui ont montré que des patients avec des lésions préfrontales unilatérales présentent une négativité dans les essais corrects de même amplitude que celle dans les erreurs. Contrairement aux patients schizophrènes, la  $N_E$  des erreurs n'est pas réduite chez les patients lésés. C'est

l'amplitude de la  $N_E$  des corrects qui est considérablement augmentée jusqu'à atteindre le même niveau que celle des erreurs.

### 1.5.3 Hyperfonctionnement des processus de contrôle

Le Trouble Obsessionnel-Compulsif ou TOC ("*Obsessive Compulsive Disorder*" en anglais) est un trouble anxieux qui se caractérise par la présence d'obsessions et/ou de compulsions qui reviennent sans arrêt. Ces rituels permettent au patient de limiter l'impact de l'anxiété générée par l'obsession. L'hyperactivité du CCA chez les patients atteints de TOC est généralement interprétée comme étant le reflet d'une anxiété importante.

Gehring *et al.* (2000) ont étudié la  $N_E$  des erreurs chez des patients TOC et ont montré qu'elle est toujours plus ample chez les patients TOC que chez les sujets sains. Plus précisément, la taille de la  $N_E$  est corrélée avec la sévérité du symptôme. Ces résultats ont été confirmés par Hajcak et Simons (2002) qui ont mis en évidence une augmentation de la  $N_E$  des corrects. Dans cette lignée, Endrass *et al.* (2008), après avoir reproduit cet effet, ont montré que la  $N_E$  des ébauches d'erreur était également plus ample chez ces patients que chez les sujets contrôles. Il y aurait donc chez ces patients un contrôle de l'action excessif qui serait présent quelle que soit la qualité de la réponse. Ces résultats électrophysiologiques ont été confirmés par des études en IRMf. Ursu *et al.* (2003), en réutilisant la tâche de Carter *et al.* (1998), ont comparé l'activation du CCA dans les erreurs et les corrects chez les patients TOC et les sujets sains. L'activation du CCA dans les erreurs et les corrects est toujours plus importante chez les patients que chez les sujets contrôles.

L'interprétation de ces résultats est délicate puisqu'il existe un nombre assez important de traits psychologiques associés à cette pathologie. Nous pouvons avancer que les patients TOC présentent un hyper-fonctionnement des processus de contrôle. Il est possible d'expliquer ces résultats avec les différentes hypothèses formulées sur la  $N_E$ . L'hypothèse selon laquelle la  $N_E$  reflète le niveau d'anxiété peut rendre compte de l'augmentation de la  $N_E$  dans les trois types d'essais. L'hypothèse alternative implique les ganglions de la base. Un dysfonctionnement au niveau du circuit fronto-stiatal-thalamique-frontal pourrait induire des boucles de feedback positifs qui pourraient expliquer la persistance excessive des activités.

### 1.5.4 Les effets des manipulations expérimentales

De façon générale et assez intuitive, la consommation d'alcool compromet les performances des sujets dans les tâches de TR. Ridderinkhof *et al.* (2002) ont montré que l'alcool induit chez les sujets

un ralentissement global des temps de réaction, une diminution des ajustements comportementaux suite à une erreur, et au niveau électrophysiologique, une réduction de la  $N_E$  dans les erreurs. Pour les auteurs, la consommation d'alcool entraînerait une surestimation des performances due à la difficulté de détection des erreurs, qui expliquerait l'absence d'ajustements post-erreur. Deux explications sont proposées : soit l'alcool induit une diminution globale et constante des processus de contrôle pendant la durée de l'intoxication, soit l'alcool perturbe ponctuellement les processus de contrôle pendant de courtes périodes, pendant lesquelles le CCA ne fonctionnerait plus correctement (à la manière des "*trous de mémoire*").

L'effet de la fatigue a été étudié par Murphy *et al.* (2006). Ils ont démontré que le maintien en éveil forcé (20h) n'influence pas l'amplitude de la  $N_E$  dans les erreurs mais agit sur l'amplitude de la  $P_E$  ainsi que sur les ajustements post-erreurs. En éveil forcé, les aspects liés à la détection de l'erreur semblent être intacts. Toutefois, les auteurs ont noté des difficultés dans la mise en place des ajustements comportementaux suite à une erreur.

Alors que la consommation d'alcool détériore la capacité à reconnaître les erreurs, la fatigue détériore la capacité à mettre en place des comportements pour compenser l'erreur. La combinaison des deux notamment dans la conduite de véhicules apparaît comme être un comportement à risque.

### 1.5.5 Bilan

Nous avons vu que les processus de contrôle se mettent en place assez tardivement chez l'Homme. Déjà présents dans l'enfance, leur efficacité augmente au cours de l'adolescence. Cette mise en place progressive des processus de contrôle serait liée à la maturation tardive du CCA ou bien à son implication tardive dans ces processus. Nous pouvons imaginer également que le CCA n'a pas toutes les informations nécessaires pour évaluer la situation dans laquelle se trouve le sujet. Le vieillissement normal induit principalement un effet de ralentissement des temps de réaction, le taux de réussite reste équivalent à celui des sujets jeunes. Une baisse de la négativité liée aux erreurs comme celle liée aux réponses correctes est observée et interprétée comme étant le reflet d'une diminution des récepteurs dopaminergiques.

Les résultats des différentes études issus de la revue de littérature sont présentés sous forme d'un tableau résumé à la figure 1.17. Nous observons, à travers le tableau, les liens entre les différents marqueurs du contrôle cognitif à travers les différentes populations de sujets. Deux aspects ressortent clairement de ce tableau. D'une part, les variations d'amplitude de la  $N_E$  dans les erreurs et les essais corrects ne sont pas directement liées à la présence ou non d'un ralentissement post-erreur.

Nous ne retrouvons pas une corrélation directe entre l'amplitude de la  $N_E$  dans les erreurs et le ralentissement post-erreur. L'étude de Mathalon *et al.* (2002) montre une diminution de l'amplitude de la  $N_E$  chez les schizophrènes par rapport aux sujets contrôles et un ralentissement post-erreur normal. En revanche, Murphy *et al.* (2006) n'ont pas observé de modulation de l'amplitude de la  $N_E$  par la fatigue mais ont montré une diminution du ralentissement post-erreur. D'autre part, la  $P_E$  semble être davantage reliée à la mise en place de l'ajustement post-erreur. Lorsque l'amplitude de la  $P_E$  est diminuée, on observe une diminution voire une absence du ralentissement post-erreur (Vlamings *et al.*, 2008; Murphy *et al.*, 2006). La  $P_E$  identifiée comme liée à la prise de conscience de l'erreur semble effectivement bien liée au ralentissement post-erreur (Nieuwenhuis *et al.*, 2001).

Malgré le nombre très important d'études rapportant des effets sur l'amplitude de la  $N_E$ , son rôle n'est pas encore clairement identifié. Dans les deux premiers chapitres expérimentaux de cette thèse, nous avons cherché à préciser son rôle.

	Auteurs	Comportement			Électrophysiologie			
		Temps de réaction	Taux de réussite	Ajustements comportementaux	Différence $N_E / N_C$	$N_E$	$N_C$	$P_E$
Enfants de 10 ans / jeunes adultes	Santesso et al. (2006) Ladouceur et al. (2007)	plus longs	diminué	Ralenissement post-erreur	$N_E > N_C$	$\nearrow$	$\nearrow$	$=$
Sujets âgés	Mathalon et al. (2003)	plus longs	équivalents	?	$N_E > N_C$	$\nearrow$	$\nearrow$	$=$
Lésions PFC unilatérale	Gehring and Knight (2002)	plus longs	équivalents	Ralenissement post-erreur	$N_E = N_C$	$=$	$\nearrow$	
Patients Alzheimer	Mathalon et al. (2003)	bcp plus longs	diminué	?	$N_E = N_C$	$\nearrow$	$\nearrow$	$\nearrow$
Enfants autistes	Vlamings et al. (2008)	équivalents	équivalent	Absents	$N_E = N_C$	$\nearrow$	$=$	$\nearrow$
Schizophrènes	Vlamings et al. (2008)	plus longs	équivalent	Ralenissement post-erreur		$\nearrow$	$\nearrow$	$=$
TOC	Endrass et al. (2008)	équivalents	équivalent	Ralenissement post-erreur	$N_E > N_C$	$\nearrow$	$\nearrow$	$=$
Consommation d'alcool	Ridderinkhof et al. (2002)	plus longs	diminué	Absents	$N_E > N_C$	$\nearrow$		
Fatigue	Murphy et al. (2006)		équivalent	Ralenissement post-erreur moins efficace	$N_E > N_C$	$=$		$\nearrow$

FIGURE 1.17 – Tableau résumé des marqueurs du contrôle cognitif dans les différentes pathologies. L'abréviation  $N_C$  est utilisée pour vouloir dire négativité dans les essais corrects.

# Chapitre 2

## Modèles

### 2.1 Introduction aux Réseaux de Neurones

Les premiers travaux sur les réseaux de neurones ont été conduits par McCulloch et Pitts (1943). Ils ont proposé une formulation théorique de l'activité neuronale, un modèle simplifié du neurone biologique : le neurone formel. Les neurones sont organisés en réseau. Les sorties des neurones d'une couche donnée vont constituer les entrées des neurones de la couche suivante (figure 2.1). Grâce à ses "dendrites", le neurone formel reçoit des influences excitatrices et inhibitrices en provenance des neurones en amont dans le réseau. Toutefois, tous les neurones en amont n'influencent pas de la même manière le neurone en aval. Cette influence dépend du poids synaptique (ou du gain synaptique) de la synapse qui est à la jonction entre l'axone du neurone en amont et la dendrite du neurone en aval. Le "corps cellulaire" du neurone en aval fait deux opérations mathématiques : il effectue la somme pondérée de ses entrées (en provenance des dendrites) puis calcule une valeur d'excitation du neurone grâce à une fonction mathématique que l'on appelle fonction de transfert  $\varphi$ . Le plus souvent, il s'agit d'une fonction à seuil (ou fonction de Heaviside). Si le seuil est dépassé, le neurone renvoie une valeur d'excitation via son "axone" aux neurones en aval dans le réseau (pour plus de détails, voir figure 2.2). McCulloch et Pitts (1943) ont montré que, théoriquement, les réseaux de neurones formels pouvaient résoudre des opérations logiques et arithmétiques complexes.

Les réseaux de neurones sont capables d'apprentissage grâce à l'utilisation de règles de modification des poids synaptiques. Plusieurs méthodes sont utilisées :

1. La première méthode proposée consiste à modifier les poids synaptiques entre les neurones en fonction de l'écart qui existe entre la réponse donnée par le neurone et la réponse désirée

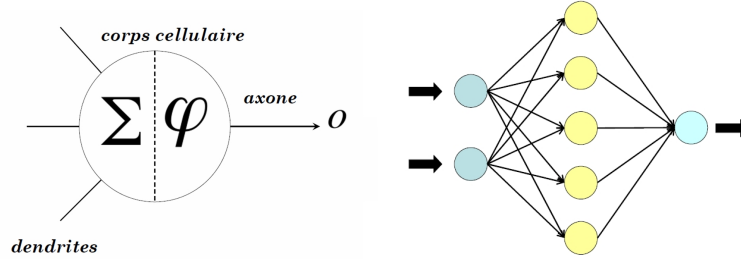


FIGURE 2.1 – **Schéma simplifié d'un neurone formel au sein d'un réseau de neurones.** Les “dendrites” du neurone reçoivent les informations en provenance des neurones en amont dans le réseau, le “corps cellulaire” effectue la somme pondérée des entrées et calcule une sortie du neurone grâce à la fonction de transfert  $\varphi$ . L’“axone” transmet la valeur d'activation du neurone aux neurones en aval du réseau. Le réseau illustré sur la partie droite de la figure est un réseau à propagation avant (“feed-forward”).

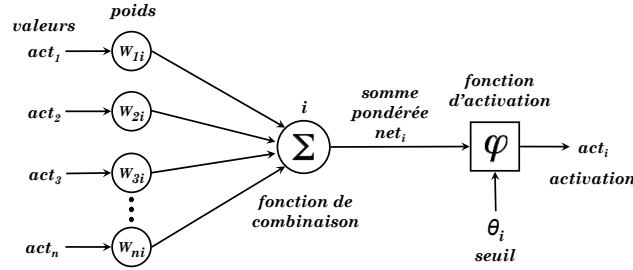


FIGURE 2.2 – **Schéma détaillé d'un neurone formel.** On note  $act_j$ , pour  $j$  variant entre 1 et  $n$ , les  $n$  informations (excitatrices ou inhibitrices) parvenant au neurone  $i$  en provenance des neurones en amont. L'influence de chacune de ces informations n'est pas équivalente, elle dépend de la valeur du poids synaptique ( $w_{ji}$ ) qui relie les neurones. Le neurone  $i$  fait la somme pondérée de ses entrées ( $net_i$ ) selon la fonction de combinaison qui a été implémentée :  $net_i = \sum_{j=1}^n w_{ji} \times act_j$ . Cette valeur est transmise à la fonction de transfert  $\varphi$  (ou fonction d'activation) qui calcule l'activation du neurone en sortie ( $act_i$ ) :  $\varphi(net_i) = act_i$ . Le plus souvent, il s'agit d'une fonction à seuil : il existe une valeur  $\theta_i$  telle que si  $net_i < \theta_i$ , alors l'activation  $act_i$  est nulle. En revanche, si  $net_i > \theta_i$ , alors le neurone est activé, ce qui se traduit par  $act_i > 0$ .



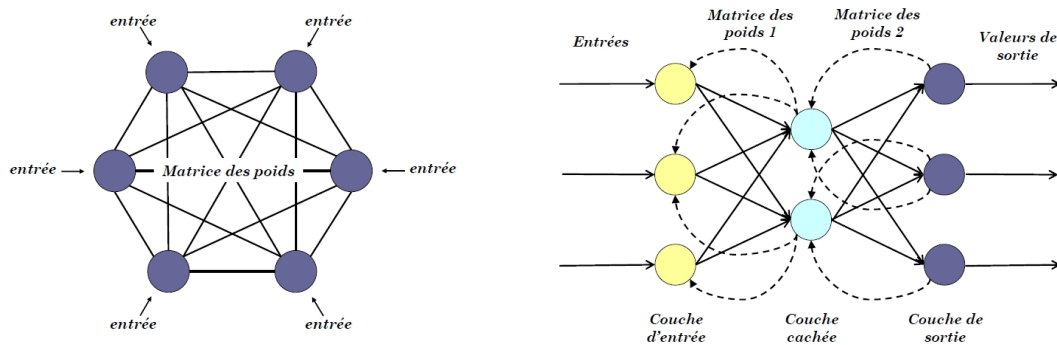


FIGURE 2.3 – **Exemples de réseaux de neurones.** À gauche, illustration d’un réseau de Hopfield ; à droite illustration d’un réseau de neurones multi-couches avec “back-propagation”

(Rosenblatt, 1958).

2. Hopfield (1982) introduit les réseaux de neurones récurrents, dits “feedback”, pour la modélisation des processus mnésiques. Ils sont composés de neurones entièrement interconnectés (figure 2.3 de gauche). La règle d’apprentissage utilisée est la règle de Hebb qui minimise l’énergie dans le réseau. Cette règle consiste à augmenter le poids synaptique de la synapse entre deux neurones lorsqu’ils sont actifs au même moment et de façon répétée (Hebb, 1949). Lorsque deux neurones adjacents sont actifs à des moments différents, alors le poids de la synapse qui les relie doit diminuer. Cette règle a été confirmée par Bliss et Lømo (1973) qui ont découvert l’existence du mécanisme de gain synaptique au niveau de l’hippocampe (la potenlisation à long terme). Les travaux de Hopfield sont une grande avancée puisqu’ils ont ouvert le champ des connexions feedback.
3. Inspirés des travaux de Hopfield, Rumelhart *et al.* (1986a,b) ont conçu un algorithme de rétropropagation du gradient de l’erreur (ou “back-propagation”), qui offre un mécanisme d’apprentissage pour les réseaux à plusieurs couches (figure 2.3 de droite). C’est une méthode qui consiste à modifier les poids synaptiques des connexions entre neurones ayant contribué à la production de l’erreur. L’ampleur de la modification des poids dépend directement de la force de l’erreur.

Toutes ces avancées dans le domaine de l’intelligence artificielle ont été utilisées par les neuroscientifiques en vue de comprendre le fonctionnement du cerveau par une modélisation des fonctions cérébrales. Pour les neuroscientifiques, l’utilisation des réseaux de neurones est intéressante à deux points de vue. D’une part, la modélisation d’une partie du fonctionnement du cerveau oblige les chercheurs à expliciter leurs hypothèses. Cette contrainte très importante participe à la difficulté

de conception des réseaux neuromimétiques. D'autre part, l'utilisation de ces réseaux de neurones offre un moyen différent et robuste de tester des hypothèses sur le fonctionnement cérébral. En effet, la traduction des hypothèses physiologiques dans un réseau de neurones permet alors d'étudier le "comportement" du réseau dans des situations particulières, et de le comparer aux comportements des individus dans une situation analogue. Les hypothèses physiologiques implémentées dans un réseau de neurones sont ainsi directement testables, ce qui les rend perfectibles mais aussi plus vulnérables. Les limites des théories sont plus facilement identifiables et permettent d'envisager des améliorations.

Concernant la modélisation des systèmes de contrôle chez l'Homme, un certain nombre de modèles neuronaux ont été proposés. Ils sont tous construits sur les bases des réseaux de neurones présentés précédemment. Dans une première partie, nous exposerons les modèles cherchant à expliquer le contrôle cognitif via les structures corticales et dans une deuxième partie les réseaux impliquant les ganglions de la base.

## 2.2 Modèles modélisant l'activité du CCA

Nous avons vu au paragraphe 1.2.1 page 16 que le cortex cingulaire antérieur est très largement impliqué dans le contrôle cognitif des actions. Observé grâce des techniques différentes aussi bien chez l'homme que chez le singe (IRMf, EEG/MEG chez l'homme ; enregistrements unitaires et enregistrements des potentiels de champs locaux chez le singe), il a souvent été proposé que le CCA ait un rôle clé dans les mécanismes physiologiques du contrôle cognitif. Nous allons décrire quatre théories sur le fonctionnement des processus de contrôle. La théorie de la "Mismatch" ainsi que la théorie émotionnelle n'ont pas fait l'objet d'une modélisation neuromimétique, les rendant plus difficilement testables. Nous décrirons beaucoup plus en détails le modèle du conflit et le modèle WEAVER ++, qui eux ont fait l'objet d'une modélisation neuromimétique. Le premier a été inspiré du modèle de Hopfield à minimisation d'énergie. Le second est issu de la psycholinguistique.

### 2.2.1 Hypothèse de la Mismatch

La première interprétation de la  $N_E$  a été qu'elle reflétait la détection de l'erreur (Falkenstein *et al.*, 1991; Gehring *et al.*, 1993). Cette hypothèse est liée au fait que la  $N_E$  a été exclusivement observée lorsque les sujets commettaient des erreurs. Cette hypothèse, appelée "mismatch" (ou "discordance"), conçoit le système de contrôle de l'action en deux parties, la première partie est chargée de détecter les erreurs, la deuxième partie est chargée de les rattraper (figure 2.4). Dans

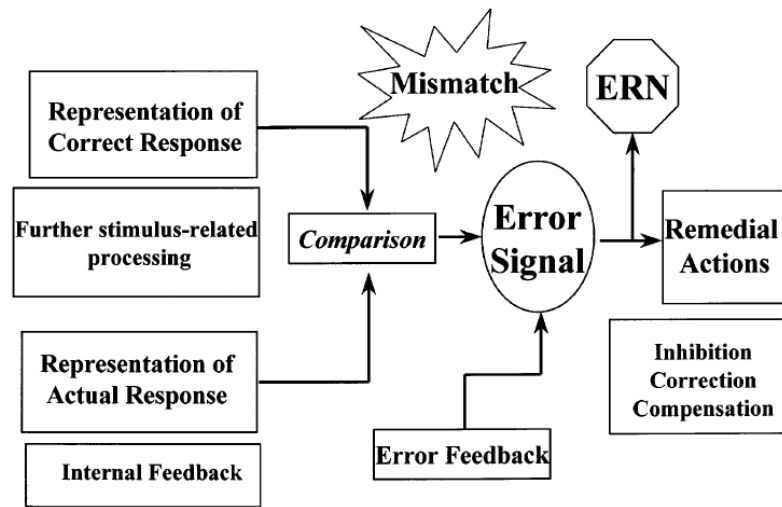


FIGURE 2.4 – Schéma représentant l'hypothèse de la Mismatch. Figure tirée de Coles *et al.* (2001).

la première partie, le système compare la représentation de la réponse correcte et la représentation de la réponse en cours. S'il y a discordance entre les deux représentations, alors un signal d'erreur est émis. Le corrélat électrophysiologique de ce signal d'erreur est la  $N_E$ . Son amplitude reflète le degré de discordance entre les deux représentations. Le signal d'erreur informe le système de rattrapage, qui met en place des actions de compensation pour l'essai en cours (inhibition de la réponse erronée ou correction de l'erreur) et des mécanismes de prévention de l'erreur pour les essais à venir.

Cette théorie n'a jamais fait l'objet d'une modélisation neuromimétique puisque l'ensemble des hypothèses qui en découlent ne sont pas explicitées. L'élément manquant à cette théorie est le fait qu'elle n'explique pas comment le système connaît la bonne réponse. Or connaître la bonne réponse est nécessaire à la comparaison entre la représentation de la réponse en cours à la représentation de la réponse correcte.

### 2.2.2 Hypothèse Émotionnelle

Avancée par de nombreux auteurs, cette théorie propose que la  $N_E$  serait liée au traitement émotionnel de l'erreur (Luu *et al.*, 2000a; Gehring et Knight, 2000; Vidal *et al.*, 2000; Gehring et Fencsik, 2001; Vocat *et al.*, 2008).

En étudiant l'impact des affects négatifs sur l'amplitude de la  $N_E$ , Luu *et al.* (2000a) ont montré que cette  $N_E$  est plus grande chez les sujets présentant davantage d'affects négatifs. Ils

ont également établis une relation entre le niveau d'engagement dans une tâche et l'amplitude de la  $N_E$ . Plus l'implication du sujet est grande, plus le fait de commettre une erreur est grave pour le sujet, et plus l'amplitude de la  $N_E$  est importante. Cette interprétation est compatible avec les résultats observés par Gehring *et al.* (1993) qui montrent une  $N_E$  plus grande dans les erreurs lorsque la consigne est de privilégier la précision par rapport à la vitesse que lorsque la consigne est de privilégier la vitesse par rapport à la précision (voir figure 1.15 page 31). Une corrélation entre le niveau d'anxiété et l'amplitude de la  $N_E$  a également été montrée (Vocat *et al.*, 2008). Hajcak *et al.* (2004) ont établi un lien entre l'importance des affects négatifs chez les sujets, la taille de la réponse électrodermale (RED) et l'amplitude de la  $N_E$  dans les essais corrects et les erreurs. Ces trois paramètres évoluent dans le même sens. En admettant que la  $N_E$  provienne du CCA, les auteurs ont avancé l'hypothèse que les affects négatifs seraient associés à une hyperactivité du CCA. Donc l'augmentation de la  $N_E$  des corrects et des erreurs pourrait être interprétée en termes de niveau d'engagement dans le système de contrôle des actions. Cette interprétation est soutenue par les données chez les patients souffrants de troubles obsessionnels compulsifs (TOC) qui ont comme caractéristique d'être particulièrement anxieux et d'avoir des difficultés à se désengager de la tâche dans laquelle ils se trouvent (Hajcak et Simons, 2002; Endrass *et al.*, 2008). Des données IRMf ont montré une activation plus importante du CCA au moment de l'erreur chez ces patients que chez les sujets contrôles (Ursu *et al.*, 2003; Fitzgerald *et al.*, 2005).

Considérer la  $N_E$  comme reflétant le traitement émotionnel de l'erreur implique que la détection de l'erreur s'est produite avant le début de la  $N_E$ . Comme pour l'hypothèse de la Mismatch, il manque une explication de la détection des erreurs par le système ce qui rend cette théorie difficilement implémentable dans un modèle neuromimétique (Yeung, 2004).

### 2.2.3 Le modèle du conflit

Le modèle de "surveillance du conflit" (par la suite dénommé "modèle du conflit") propose une explication simple et élégante pour rendre compte, à la fois, des données neurophysiologiques et comportementales liées au contrôle (Botvinick *et al.*, 2001, 2004; Yeung *et al.*, 2004; Carter *et al.*, 1998).

Le modèle du conflit se greffe sur des modélisations déjà existantes de tâches dans lesquelles plusieurs réponses sont en compétition. Le plus souvent, il s'agit de tâches manipulant la compatibilité Stimulus-Réponse. Cette modélisation, que l'on appelle "module tâche", est totalement autonome et doit rendre compte des différentes étapes perceptives et motrices du traitement de

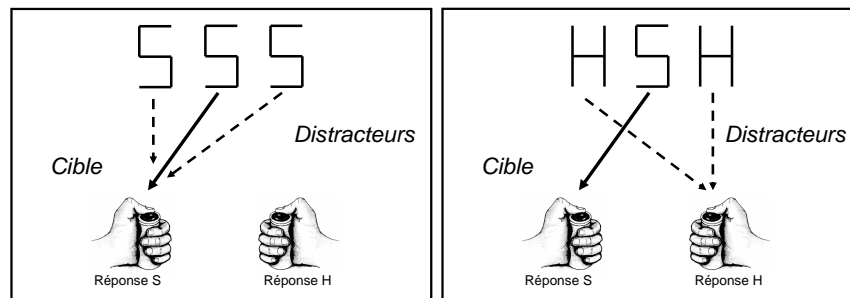


FIGURE 2.5 – **Réalisation de la tâche d'Eriksen (Eriksen et Eriksen, 1974).** Dans cette tâche, les stimuli sont composés de trois lettres. On demande aux sujets d'identifier la lettre centrale et d'ignorer les deux autres lettres. Lorsque les trois lettres sont identiques, comme sur l'illustration de gauche, l'essai est dit "compatible". La cible ainsi que les distracteurs "appellent" la même réponse. Dans cet exemple, la consigne est d'appuyer sur le bonton gauche lorsque la lettre centrale est un *S* et sur le bonton droit lorsque la lettre cible est un *H*. Sur l'illustration de droite, l'essai est dit "incompatible". Cela implique que la cible et les distracteurs appellent des réponses opposées.

l'information spécifiques à cette tâche. Il s'agit d'une modélisation neuromimétique composé au moins de deux couches : une couche des stimuli et une couche de réponse chargées respectivement de la "perception" des stimuli et de la "production" de la réponse. Nous prendrons en exemple la tâche développée par Eriksen et Eriksen (1974) (voir figure 2.5).

Le modèle du conflit est composé de deux parties. La première partie, appelée "module du conflit", est greffée sur le module tâche au niveau de la couche de réponse et mesure la dynamique d'activations des unités de réponse tout au long de l'essai en calculant le degré de co-activation des réponses en compétition pendant toute la durée de l'essai. La deuxième partie du modèle du conflit est le "module attentionnel" greffé au niveau des unités perceptives de la couche des stimuli. Lorsque celui-ci est sollicité par le module du conflit, ce module attentionnel exerce une influence sur les unités de la couche des stimuli du module tâche dans le but de biaiser le traitement des prochains stimuli. La figure 2.6 illustre le modèle du conflit greffé sur une modélisation de la tâche d'Eriksen et Eriksen (1974).

### Les grandes hypothèses

1. Un module du réseau mesure le conflit, c'est-à-dire le degré de compétition entre les réponses pendant toute la durée de l'essai (voir plus bas pour une description plus formelle).
2. À la fin de chaque essai, le module du conflit informe le module attentionnel de la quantité de conflit enregistré pendant toute la durée de l'essai.

3. Le module attentionnel, alerté par le module du conflit, est chargé de la mise en place des ajustements comportementaux afin de minimiser le risque d'erreur lors des prochains essais.

Il permet une refocalisation attentionnelle sur les caractéristiques pertinentes de la tâche.

Les données expérimentales et les résultats des simulations ont conduit les auteurs à réinterpréter l'activité du CCA observée en IRMf et de la  $N_E$  observée en EEG comme étant la manifestation neurophysiologique du conflit (Botvinick *et al.*, 2001, 2004; Yeung *et al.*, 2004; Carter *et al.*, 1998). L'activité du module attentionnel est attribuée à l'activité du cortex préfrontal (MacDonald *et al.*, 2000; Kerns *et al.*, 2004). L'hypothèse de refocalisation attentionnelle a été testée par Egner et Hirsch (2005a). L'avancée majeure du modèle du conflit par rapport aux précédents modèles, est qu'il est capable d'expliquer comment le contrôle "sait" qu'il doit entrer en jeu.

Dans un premier temps, nous allons décrire la tâche d'Eriksen et la modélisation neuromimétique de cette tâche. Puis nous nous attacherons à décrire comment est greffé le modèle du conflit sur ce modèle neuromimétique.

### Description et modélisation de la Tâche d'Eriksen

Dans la tâche d'Eriksen et Eriksen (1974), les stimuli sont un ensemble de trois lettres (S ou H) présentées côtes à côtes. La consigne du sujet est d'identifier la lettre centrale et de donner la réponse associée. La première et la troisième lettre sont toujours identiques et jouent le rôle de distracteurs. Nous appelons "Compatibles" les stimuli dont la cible (lettre centrale) et les distracteurs (1ère et 3ème lettre) sont identiques ( $SSS$  et  $HHH$ ); et "Incompatibles" les stimuli dont la cible et les distracteurs sont différents ( $SHS$  et  $HSH$ ). Un exemple de réalisation de la tâche d'Eriksen est illustré dans la figure 2.5. Dans cet exemple, le sujet doit appuyer sur le bouton de gauche lorsque la lettre centrale est un "S", et sur le bouton de droite lorsque la lettre centrale est un "H".

Le modèle de la tâche d'Eriksen est composé de deux couches : la couche des stimuli et la couche des réponses (partie rose de la figure 2.6). La couche des stimuli est chargée de coder l'ensemble des attributs des stimuli. Puisque que chacune des trois lettres peut être un "S" ou un "H", deux unités sont nécessaires pour coder le distracteur de gauche ( $H_g$  et  $S_g$ ), deux autres pour coder la lettre cible ( $H_c$  et  $S_c$ ) et enfin deux pour coder le distracteur de droite ( $H_d$  et  $S_d$ ). Chaque unité de la couche des stimuli est reliée à une unité de la couche des réponses. Dans la couche des réponses, le nombre d'unités de réponse correspond au nombre de réponses possibles. Dans le cas de la tâche d'Eriksen, seulement deux réponses sont possibles ("S" et "H").

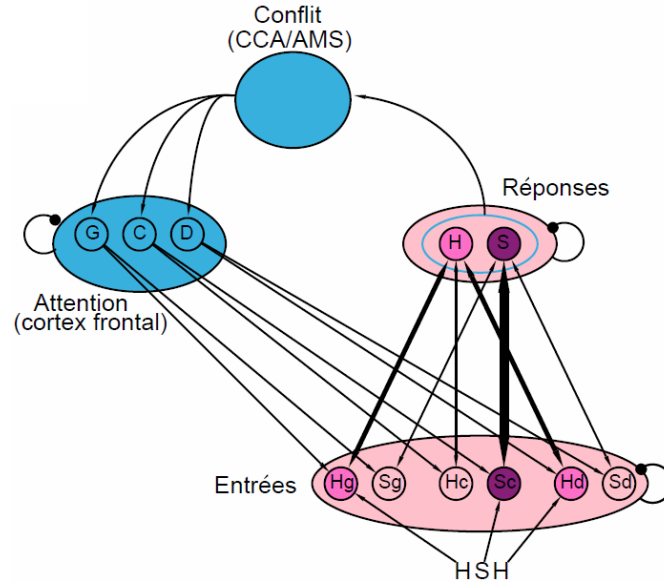


FIGURE 2.6 – **Modèle du Conflit de Botvinick *et al.* (2001) adapté pour la tâche d'Eriksen (Eriksen et Eriksen, 1974).**

La modélisation de la tâche d'Eriksen est composée de deux couches : une couche des stimuli et une couche de réponse (en rose). Parmi les six unités qui composent la couche des stimuli, seulement trois sont activées par le stimulus. Les unités *Sc* et *Hc* symbolisent la lettre centrale et les unités *Sg*, *Hg*, *Sd* et *Hd* symbolisent les distracteurs situés à gauche et à droite de la lettre cible. Chacune de ces unités active l'une des deux unités de la couche de réponse (unités *H* ou *S*). Il existe une compétition entre ces unités. Dès que l'une d'elles atteint le seuil d'activation de la réponse, la réponse est donnée.

Le modèle du conflit (en bleu) est composé de deux couches : le module du conflit et la couche attentionnelle. Le module du conflit enregistre le degré de compétition entre les réponses au cours de l'essai, ce qui correspondrait à l'activité du CCA, et influencerait les aires responsables de l'allocation de ressources attentionnelles au niveau du cortex préfrontal. Ce module est relié aux unités de la couche attentionnelle : les unités *C*, *D* et *G*. L'unité *C* est chargée de remobiliser l'attention sur la lettre centrale du stimulus (unités *Hc* et *Sc* de la couche des stimuli), tandis que les unités *D* et *G* influencent les distracteurs (unités *Hd* et *Sd*, *Hg* et *Sg*). Lorsque le conflit est non-nul, l'unité *C* devient plus active alors que les unités *G* et *D* diminuent. Si le conflit est nul, l'activité de l'unité *C* décroît tandis que l'activité des unités *G* et *D* augmente.

La force des connexions entre les unités de la couche des stimuli et les unités de la couche de réponse est fonction de la nature de l'unité de la couche des stimuli. Les connexions entre les unités de la couche des stimuli chargées de traiter la cible au sein du stimulus ( $H_c$  et  $S_c$ ) ont des connexions plus fortes avec les unités de la couche des réponses (respectivement les unités  $H$  et  $S$ ) que les unités de la couche des stimuli chargées de traiter les distracteurs au sein du stimulus ( $H_g$ ,  $S_g$ ,  $H_d$  et  $S_d$ ). Dans la couche des réponses, les unités de réponse sont en compétition, elles ont une action inhibitrice l'une envers l'autre. Lorsque le niveau d'activation de l'une d'elles dépasse le seuil de réponse, alors la réponse est déclenchée.

### Déroulement d'un essai

Toutes les unités dans le réseau fonctionnent selon le modèle du neurone formel de McCulloch et Pitts (1943) (voir paragraphe 2.1 page 43, et figure 2.2 page 44). À chaque cycle, chaque unité  $i$  calcule son net input ( $net_i$ ) grâce à la fonction de combinaison implémentée dans le modèle.

Pour les unités de la couche des stimuli ( $H_g$ ,  $S_g$ ,  $H_c$ ,  $S_c$ ,  $H_d$  et  $S_d$ ) et les unités de la couche de réponse ( $H$  et  $S$ ), le  $net_i$  de l'unité  $i$  est calculé de cette manière :

$$net_i = (ext_i * estr) + \sum_{j=1}^N act_j w_{ji} + noise \quad (2.1)$$

où :  $ext_i$  correspond aux entrées extérieures de l'unité  $i$ .  $estr$  est une constante.  $noise$  correspond à du bruit aléatoire dont les valeurs sont issues d'une distribution normale (de moyenne 0 et d'écart-type fixe). L'expression  $\sum_{j=1}^N act_j w_{ji}$  correspond à la somme pondérée des entrées en provenance des unités du module tâche. Soit  $j$  variant de 1 à  $N$  avec  $N$  correspondant au nombre d'unités dans le modèle ayant une influence sur l'unité  $i$ .  $act_j$  correspond au niveau d'activité de l'unité  $j$ . Cette activité est multipliée par le poids de la connexion entre l'unité  $j$  et l'unité  $i$  ( $w_{ji}$ ). La valeur de ce poids peut être positive ou négative en fonction de la nature excitatrice ou inhibitrice de la connexion entre les neurones.

Il existe deux types d'entrées extérieures  $ext_i$  pour des unités de la couche des stimuli : (1) lors de "l'arrivée du stimulus", les unités concernées par le stimulus ont leur  $ext_i$  augmenté d'une constante ; (2) la deuxième entrée extérieure correspond à l'influence des unités de la couche attentionnelle (détaillées plus bas). Les entrées extérieures des unités de la couche de réponse correspondent à une constante qui est présente uniquement entre le début de l'essai et la présentation du stimulus. Le  $net_i$  correspond donc à la somme des entrées externes et internes. À partir du  $net_i$ , la fonction de transfert calcule la nouvelle activation de l'unité ( $act_i$ ).



Lors de la “présentation du stimulus”, les entrées externes des unités concernées par le stimulus sont augmentées. En prenant pour exemple du stimulus  $HSH$ , les unités dont le niveau des entrées est augmenté sont les unités  $H_g$ ,  $S_c$  et  $H_d$ . Par cascade, le niveau des entrées des unités  $S$  et  $H$  va être augmenté. L'information se “déplace” ainsi dans le réseau. Le modèle fournit sa réponse dès que l'une des deux unités de réponse aura franchi le seuil de réponse.

### Description du modèle du Conflit

En parallèle, le module du conflit mesure le degré de compétition entre les unités de la couche de réponse. Ce degré de compétition est défini par le degré de co-activation entre les unités de réponses pendant toute la durée de l'essai (même au-delà de la réponse) et se calcule grâce au produit de l'activation des unités de réponse (uniquement lorsqu'elles sont strictement positives) calculé à chaque pas de temps (ou cycle) suivant cette équation :

$$conflit(t) = \begin{cases} act_H(t) \times act_S(t) \times w & \text{si } act_H(t) > 0 \text{ et } act_S(t) > 0 \\ 0 & \text{pour les autres cas} \end{cases} \quad (2.2)$$

où  $act_H(t)$  et  $act_S(t)$  sont les niveaux d'activation des unités de réponse  $H$  et  $S$  à l'instant  $t$ .  $w$  correspond au poids de l'inhibition latérale entre les deux unités.

Le conflit est intégré au cours de l'essai. La mesure finale ( $Co(e)$ ) reflète la quantité de conflit présente pendant la durée ( $T$ ) de l'essai ( $e$ ), tel que :

$$Co(e) = \sum_{t=0}^T conflit(t) \quad (2.3)$$

La couche attentionnelle est composée de trois unités : l'unité  $C$ , l'unité  $G$  et l'unité  $D$ . L'unité  $C$  dite “centrale”, remobilise l'attention sur la partie centrale du stimulus à l'essai suivant, en agissant sur les unités  $Hc$  et  $Sc$  de la couche des stimuli (voir équation 2.4). Les unités  $G$  et  $D$  font varier le niveau d'attention sur la partie gauche et la partie droite du stimulus : sur les unités  $Hg$ ,  $Sg$  et les unités  $Hd$ ,  $Sd$  (voir équation 2.5).

L'activité de l'unité  $C$  ( $ext_C$ ) dépend du niveau de conflit enregistré à l'essai précédant ( $Co(e-1)$ ) ainsi que sa propre activité à l'essai précédent ( $ext_C(e-1)$ ).

$$ext_C(e) = \lambda ext_C(e-1) + (1-\lambda)(\alpha Co(e-1) + \beta) \quad (2.4)$$

où  $\alpha$ ,  $\beta$  et  $\lambda$  correspondent à trois constantes.

Le calcul de l'activité des unités  $D$  et  $G$  résulte du niveau d'activité de l'unité  $C$ .

$$ext_G = ext_D = \frac{3 - ext_C}{2} \quad (2.5)$$

Chacune des unités de la couche des stimuli est reliée à l'une des trois unités de la couche attentionnelle. Les variables  $ext_G$ ,  $ext_C$  et  $ext_D$  participent, respectivement, aux entrées externes des unités de la couche des stimuli  $H_g$  et  $S_g$ ,  $H_c$  et  $S_c$  et  $H_d$  et  $S_d$  (voir équation 2.1).

### Dynamique du modèle

Ce modèle vise à rendre compte, entre autre, des ajustements comportementaux (effets séquentiels et ralentissement post-erreurs, voir paragraphes 1.1.2 et 1.1.3 page 12) ainsi que des changements d'activité du CCA et du cortex préfrontal enregistrés en IRMf et de la dynamique temporelle de l'onde d'erreur dans les erreurs ainsi que dans les essais corrects.

Si le paramètre aléatoire *noise* est enlevé de l'équation permettant de calculer l'activation des unités à chaque cycle (équation 2.1) et en l'absence du module du conflit (sans entrées extérieures en provenance de la couche attentionnelle), le module tâche est conçu de façon à ce que la réponse correcte soit fournie à chaque essai quelle que soit la nature du stimulus (compatible ou incompatible). Le module du conflit et le paramètre aléatoire *noise* permettent au modèle de pouvoir rendre compte des variations de performance observées chez les sujets humains.

Prenons l'exemple d'un stimulus compatible *SSS*, les trois unités de la couche des stimuli, dont le niveau d'activation a été augmenté suite à l'apparition du stimulus, activent la même réponse *S*. La somme pondérée des entrées de l'unité de réponse sollicitée va être largement augmentée dans les prochains cycles, et le seuil de déclenchement de la réponse sera franchi. La réponse alternative n'ayant pas été stimulée (réponse *H*), il est probable qu'à aucun moment au cours de la durée de l'essai, les deux réponses aient été activées simultanément. Si le conflit est nul (absence de co-activation des réponses, équations 3.2 et 2.3), l'activité de l'unité *C* de la couche attentionnelle va diminuer (voir équation 2.4) et par conséquent l'activité des unités *G* et *D* va être augmentée (voir équation 2.5). Ce qui implique qu'au prochain essai, les entrées extérieures des unités chargées du traitement des distracteurs seront un peu augmentées et les entrées extérieures des unités chargées du traitement de la cible seront diminuées.

Lorsque l'essai est incompatible, les unités de la couche des stimuli sollicitées par le stimulus ont leur niveau d'activation augmenté. La cible et les distracteurs étant différents, les deux unités de réponse vont être activées. Lorsque l'une d'elles atteint le seuil de déclenchement de la réponse, la réponse est donnée. Comme nous l'avons déjà dit, le module tâche est programmé pour converger dans tous les cas vers la bonne réponse. La question est : le niveau d'activation de la réponse incorrecte va-t-il franchir le seuil de réponse avant l'unité de réponse correcte ? La figure 2.7 illustre les deux cas de figure. Dans la figure de droite, la bonne réponse a été donnée. Le maximum

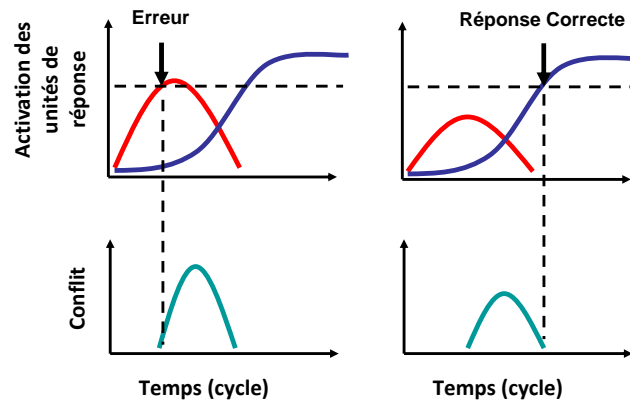


FIGURE 2.7 – **Simulation d’une erreur et d’un essai correct.** Les courbes bleues et rouges correspondent respectivement à l’activation de l’unité de réponse correcte et à l’activation de l’unité de réponse incorrecte en fonction des cycles. Le trait horizontal en pointillés correspond au seuil de réponse. Le niveau de conflit, représenté en vert, est calculé à chaque cycle. Dans le cas des erreurs, le conflit est observé après la réponse et dans le cas des essais corrects le conflit est observé après la réponse.

d’activité du conflit enregistré est avant la réponse. Dans le cas de figure où l’erreur a été donnée, le conflit est maximal après le déclenchement de la réponse (figure 2.7 de droite). Dans les deux cas de figure, un conflit ayant été enregistré, le niveau d’activation de l’unité  $C$  de la couche attentionnelle va augmenter proportionnellement en fonction de la quantité de conflit, tandis que le niveau d’activation des unités  $G$  et  $D$  va diminuer (voir équations 2.4 et 2.5). Au prochain essai, l’influence de l’attribut pertinent sera augmentée et l’influence de l’attribut non-pertinent sera diminuée.

### Modèle du Conflit et marqueurs neurophysiologiques du contrôle

Dans la comparaison compatible-incompatible, des différences physiologiques ont été observées en IRMf et EEG. En IRMf, l’activité dans le CCA est plus importante en condition incompatible et le cortex préfrontal est plus activé suite à un essai incompatible (voir paragraphes 1.2.1). La  $N_E$  est présente après la réponse dans le cas des erreurs dans les essais compatibles et incompatibles. Dans le cas des essais corrects incompatibles, une onde similaire, la  $N_2$ , est présente avant la réponse. Son amplitude est plus petite que celle de la  $N_E$  enregistrée dans les erreurs (voir paragraphes 1.3.8).

Les simulations du modèle du conflit réalisées par Botvinick *et al.* (2001) permettent de rendre compte des données expérimentales. Au niveau comportemental, l’effet de compatibilité, les effets

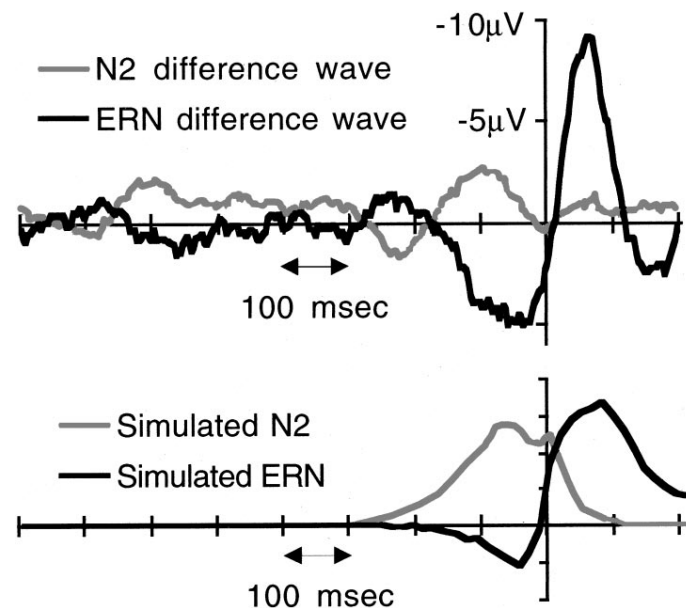


FIGURE 2.8 – Les activités  $N_E$  et  $N2$  observées et comparées aux simulations du modèle du conflit par Yeung *et al.* (2004). En haut, il s'agit des données expérimentales obtenus en EEG classique (monopolaire) au niveau l'électrode FCz chez des sujets sains pendant qu'ils effectuaient la tâche d'Eriksen. En bas, il s'agit des tracés correspondant de l'activité du module du conflit pendant une simulation de la tâche d'Eriksen. Dans les deux graphiques, le zéro des temps correspond au moment où la réponse a été donnée. En trait clair, il s'agit de la différence entre l'activité enregistrée dans les essais incompatibles corrects et dans les essais compatibles corrects. En trait foncé, il s'agit de la différence moyenne entre l'activité dans les erreurs et les essais corrects. Nous avons en trait foncé la  $N_E$  et la  $N_E$ -simulée qui débute juste avant la réponse et qui est à son maximum juste après la réponse. En trait clair, nous avons la  $N2$  et la  $N2$ -simulée, de plus petite amplitude, qui se développe et dont le maximum d'activité survient avant la réponse correcte.

séquentiels, l'accélération au moment de l'erreur et le ralentissement post-erreur sont reproduits dans les simulations du modèle. Au niveau électrophysiologique, la  $N_E$  dans les erreurs semble bien refléter le niveau de conflit. Yeung *et al.* (2004) ont montré que le modèle du conflit pouvait rendre compte de la  $N2$  dans les essais corrects (voir figure 2.8).

Le lien entre le module du conflit et la couche attentionnel est conforté par la mise en évidence par une étude IRMf de la corrélation entre l'activité du CCA et l'activité du cortex préfrontal (Kerns *et al.*, 2004).

L'hypothèse de refocalisation attentionnelle a été testée directement par Egner et Hirsch (2005a) en IRMf. Les auteurs ont enregistré les variations d'activité de la réponse BOLD dans une tâche construite sur le modèle de la tâche de Stroop avec des visages et des noms de personnages célèbres (figure 2.9). La tâche des sujets était de classer les personnages dans la catégorie "homme politique"



FIGURE 2.9 – **Exemples de stimuli utilisés par Egner et Hirsch (2005a) dans leur étude IRMf.** Dans les versions NOM et VISAGE de cette tâche, on demande aux sujets de classer le personnage auquel fait référence le nom ou le visage comme appartenant à la catégorie “acteur” ou “homme politique”. Les stimuli de la première ligne sont compatibles puisque le visage et le nom appartiennent à la même catégorie : des acteurs à gauche et des hommes politiques à droite. Les stimuli de la deuxième ligne sont incompatibles (figure tirée de Egner et Hirsch, 2005a)

ou dans la catégorie “acteur de cinéma”. Dans la tâche VISAGE, les sujets devaient opérer cette catégorisation sur le visage et ne pas tenir compte du nom ; et dans la tâche NOM, ils devaient l’effectuer sur le nom et ne pas tenir compte du visage. Dans les essais compatibles, le nom et le visage faisaient référence à une même catégorie. Inversement, dans les essais incompatibles, le visage et le nom appartenaient à des catégories différentes. Les temps de réaction moyens montrent qu’il existe une modulation de l’effet de compatibilité à l’essai  $n$  en fonction de la nature de l’essai  $n-1$ . Dans les deux tâches, l’effet de compatibilité est réduit suite à un essai incompatible. Les auteurs se sont intéressés aux variations d’activité de l’aire spécifique du traitement des visages (FFA pour “Fusiform Face Area”) en fonction de la compatibilité à l’essai  $n$ , la compatibilité à l’essai  $n-1$ , et en fonction de la pertinence de l’information sur les visages pour la tâche (l’attribut visage est pertinent pour la tâche VISAGE, et non-pertinent pour la tâche NOM). Lorsque les sujets devaient réaliser la tâche NOM, les auteurs n’ont observé de différences dans l’activité de la FFA ni en fonction de la compatibilité à l’essai  $n$ , ni en fonction de la compatibilité à l’essai  $n-1$ . En revanche, ils ont montré une modulation de l’activité de la FFA en fonction de la compatibilité à l’essai  $n$  et à l’essai  $n-1$  dans la tâche VISAGE qui pourrait expliquer les différences de TR observées en fonction des conditions. Ces résultats indiquent que les effets séquentiels obtenus sur les TR sont explicables par une augmentation du traitement de l’attribut pertinent et non par une diminution du traitement de l’attribut non pertinent. Ces résultats, obtenus par Egner et Hirsch (2005a), viennent préciser le mode d’action du cortex préfrontal sur les structures chargées du traitement

du stimulus et appuient partiellement l'une des hypothèses fondamentales du modèle du conflit (hypothèse 3, page 49). L'augmentation du traitement de l'attribut non-pertinent après un essai incompatible semble bien être présente. La diminution du traitement de l'attribut non-pertinent n'a pas été observée alors que l'équation 2.5 du modèle du conflit prévoit cette diminution.

### Remises en cause du modèle du conflit

Dans le modèle du conflit, la couche attentionnelle, qui reflèterait l'activité du cortex préfrontal, est sollicitée seulement à la fin de l'essai. Or cette notion de "fin d'essai" ne fait pas référence à une réelle dynamique temporelle : le modèle ne propose pas de moyen permettant de savoir à quel moment l'essai est terminé et de surcroît à quel moment les ajustements comportementaux doivent être mis en place. Les auteurs du modèle du conflit proposent une explication pour rendre compte de la capacité à détecter les besoins d'ajustements comportementaux mais ne détaillent pas comment le système sait le moment où ces ajustements comportementaux doivent être amorcés.

Du fait de la mise en place des ajustements comportementaux en fin d'essai, aucune remise en cause de la commande motrice n'est envisagée en cours d'essai. Or, Allain *et al.* (2004b) ont montré l'existence d'un système superviseur opérant en ligne. Ils ont comparé la forme de la bouffée EMG des réponses correctes et des erreurs. Le début des bouffées EMG est identique pendant environ 30 ms, puis les deux courbes se dissocient. Celle des erreurs est alors diminuée par rapport à celle des essais corrects (voir figure 2.10). Ces résultats montrent que la commande motrice corticale initiale est identique entre 0 et 30 ms après le début de l'EMG. La diminution de la bouffée EMG dans les erreurs est interprétée comme étant une tentative d'inhibition en ligne de la réponse erronée.

Il a été montré que dans certains essais corrects contenant une ébauche d'erreur, une  $N_E$  apparaissait au moment de l'ébauche. Selon les hypothèses fondamentales du modèle du conflit (page 49), les essais corrects présentant une  $N_E$  devraient être suivis d'ajustements comportementaux aux essais suivants. Burle *et al.* (2005) ont testé cette prédiction et ont montré qu'il n'y avait pas de réduction de l'effet de compatibilité suite à un essai correct avec ébauche d'erreur comparé à un essai pure-correct (sans ébauche d'erreur). Ce résultat remet sérieusement en cause l'une des hypothèses majeures du modèle du conflit. Ces constats montrent que le modèle du conflit présente des imprécisions et des failles notamment sur la mise en place des ajustements comportementaux.

Chez le singe effectuant une tâche de compatibilité spatiale (version oculaire de la tâche de Simon), Nakamura *et al.* (2005) ont observé, en enregistrement unitaire, la décharge des neurones dans le champ oculaire supplémentaire (supplementary eye field) et le CCA. Ni dans le CCA, ni dans le champ oculaire supplémentaire, les auteurs n'ont pu identifier de neurones dont l'activité

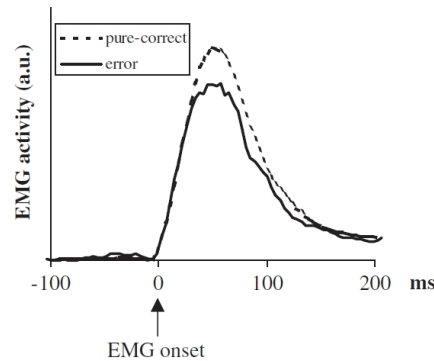


FIGURE 2.10 – **Activités électromyographiques moyennées (redressées et intégrées) et calées sur le début de l'activité musculaire dans les erreurs et les essais corrects.** Figure tirée de Allain *et al.* (2004b).

pourrait correspondre au conflit au sens de Botvinick *et al.* (2001).

Dans une tâche de STOP saccadique chez le macaque, Emeric *et al.* (2008) ont enregistré les potentiels de champ locaux (local fiels potential : LFP) au sein du CCA. Les animaux devaient exécuter une saccade vers une cible pour obtenir la récompense. Dans certains essais, un signal STOP apparaissait au niveau du point de fixation central quelques millisecondes après le signal GO, les animaux devaient stopper leur saccade pour obtenir la récompense. Dans ces essais STOP, deux programmes moteurs sont en compétition : la programmation de la saccade et l'interruption de la saccade. Ainsi, selon l'une des hypothèses du modèle du conflit, les essais STOP-corrects devraient contenir davantage de conflit que les essais classiques corrects (GO-corrects) puisque deux programmes moteurs étaient en compétition. Les auteurs ont calculé la différence d'activité dans le CCA dans les conditions GO-corrects et STOP-corrects dans la fenêtre temporelle autour du TR des essais GO-corrects. Ils ont montré que lorsque le signal STOP arrive tardivement (conflit important), la différence d'activité du CCA entre les essais STOP-corrects et GO-Corrects diminue. Le modèle du conflit n'a pas été testé dans cette tâche. Toutefois, si le raisonnement des auteurs est correct, alors ce résultat est contraire à l'idée que l'activité du CCA reflète le niveau de conflit.

Ces études chez l'homme et le singe montrent que le CCA ne semble pas être actif selon les prédictions du modèle du conflit et n'apparaît pas comme étant systématiquement relié à la mise en place des ajustements comportementaux. Dans le premier chapitre expérimental, nous avons testé l'une des hypothèses fondamentales du modèle du conflit en comparant les résultats d'une étude de simulation et d'une étude expérimentale.

### 2.2.4 Modèle Weaver ++ (Roelofs, 2003)

WEAVER ++ est un modèle computationnel créé pour rendre compte des effets sémantiques et phonologiques dans des tâches de production vocale de mots (Levelt *et al.*, 1999; Roelofs, 1992, 1997). Initialement, ce modèle n'a pas été conçu pour expliquer l'effet de compatibilité Stimulus-Réponse dans la tâche de Stroop (Stroop, 1935), mais Roelofs (2003) a ajouté un module chargé d'implémenter le contrôle cognitif afin que le modèle soit capable de rendre compte de cet effet. L'activité de ce module dans le réseau WEAVER ++ correspondrait à l'activité du cortex cingulaire antérieur.

#### Description du modèle

Ce modèle est dit “hybride” puisqu'il s'agit d'un réseau de neurones dont la propagation de l'activité est régie par un ensemble de règles. Le réseau est composé de trois couches (figure 2.11) :

- La première couche est appelée **la couche des concepts**. Chacune des unités appartenant à cette couche représentent un concept. Les connexions entre les unités dépendent du lien sémantique qui existe entre les concepts qu'elles représentent. Par exemple, le concept ROUGE(X) est, entre autres, relié aux concepts COULEUR(X) et VERT(X) (voir figure 2.11).
- La deuxième couche est appelée **la couche syntaxique**. Les unités représentent des lemmes (ou unités lexicales). Un lemme est une suite de caractères formant une unité sémantique. Les connexions entre les unités de cette couche existent lorsque les lemmes appartiennent à la même classe syntaxique (par exemple, les lemmes *rouge* et *vert* sont connectés puisqu'ils appartiennent à la classe syntaxique des adjectifs).
- La troisième couche est appelée **la couche de la forme du mot** représentant les morphèmes (<rouge>), les segments (/r/ /o/ /u/ /g/ /e/) et les programmes moteurs nécessaires à l'articulation du mot ([rouge]).

Il existe aussi des connexions entre les unités des différentes couches. Par exemple, le concept ROUGE(X) est relié au lemme *rouge* qui est lui-même relié aux unités de la couche de la forme du mot correspondant au lemme *rouge*.

#### Diffusion de l'activation dans les unités du modèle

L'activation dans le réseau se diffuse selon les lois de diffusion classiques. À chaque cycle, les unités calculent la somme pondérée des entrées qu'elles reçoivent en provenance des unités en



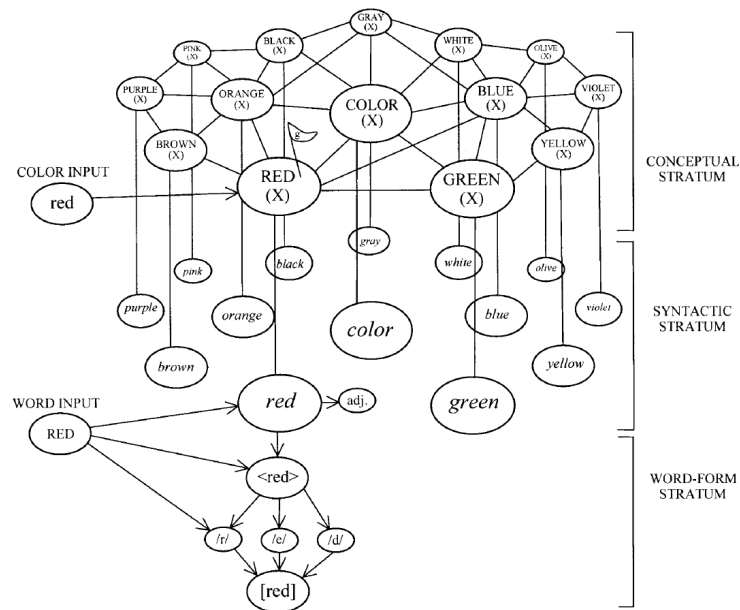


FIGURE 2.11 – **Schéma du modèle Weaver ++ adapté de Roelofs (1992) pour la tâche de Stroop.** Figure tirée de Roelofs (2003)

amont du réseau, et calculent leur propre valeur d'excitation qu'elles renvoient aux unités en aval dans le réseau (voir paragraphe 2.1 page 43, et figure 2.2 page 44). Toutefois, la propagation de l'activité dans le réseau dépend également d'un certain nombre de règles qui sont élaborées en fonction de la tâche.

### Les règles associées au modèle

Du fait de l'architecture du réseau, la distance à parcourir dans le réseau entre le lemme et le programme moteur (*rouge* et [rouge]) est plus courte que la distance à parcourir entre le concept et le programme moteur (*ROUGE*(*X*) et [rouge], voir figure 2.11). Sans la mise en place de règles, la réponse donnée serait toujours celle reliée au lemme sélectionné par la lecture du mot puisque le seuil de déclenchement de la réponse sera franchi plus rapidement pour la réponse liée à la lecture du mot que pour la réponse liée à la couleur du mot.

Le modèle *WEAVER ++* fait comme hypothèse que le cortex cingulaire antérieur (CCA) implémente le contrôle cognitif en fonction de la tâche. Il mène en parallèle deux actions. D'une part, il augmente sélectivement le traitement de l'attribut pertinent pour la tâche, et d'autre part, il bloque le traitement de l'attribut non pertinent.

Les règles de fonctionnement du modèle *WEAVER ++* entrent en vigueur dès qu'une unité

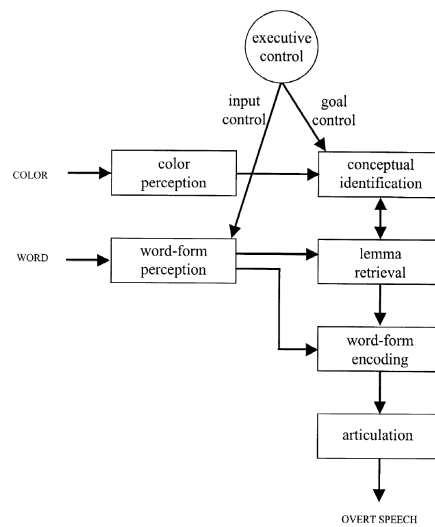


FIGURE 2.12 – Schéma du modèle Weaver ++ avec le CCA, adapté de Roelofs (1992)

concernée par une règle devient active.

**Règle 1 :** Si la tâche est de nommer la couleur du stimulus, et que le concept correspondant à la couleur du stimulus est actif, alors il faut sélectionner ce concept et le signaler comme étant le concept “cible” et augmenter son activation.

**Règle 2 :** Si le lemme relié au concept cible a atteint le seuil d’activation de la réponse, alors il faut sélectionner ce lemme et permettre son articulation vocale.

**Règle 3 :** Si la tâche est de nommer la couleur du stimulus, et que l’entrée en provenance d’un mot est active alors, il faut bloquer cette entrée.

Le contrôle se met en place suivant les règles 1 et 3. La règle 2 est une règle générale permettant de sélectionner la réponse.

### Déroulement d’un essai

Prenons en exemple un stimulus incompatible : **VERT**. La lecture du mot active le lemme *vert* et la couleur du mot active le concept ROUGE(X). Le système du contrôle exécutif se met en place en utilisant les règles 1 et 3. Suivant la règle 1, l’activité du concept ROUGE(X) est identifiée comme étant la cible et son niveau d’activation est augmenté. Suivant la règle 3, l’activité du lemme *vert* va être bloquée. Le concept ROUGE(X) stimule le niveau d’activation du lemme *rouge*, et celui-ci va franchir le seuil de déclenchement de la réponse (règle 2), le programme moteur [rouge] sera exécuté.

Prenons en exemple un stimulus compatible : **BLEU**. La lecture du mot active le lemme *bleu* et la couleur du mot active le concept BLEU(X). Le système du contrôle exécutif se met en place en utilisant uniquement la règle 1. Suivant cette règle, l'activité du concept BLEU(X) est identifiée comme étant la cible et son niveau d'activation est augmenté. Le réseau identifie alors le stimulus comme étant compatible. Il n'y a pas de blocage du lemme *bleu* puisque celui-ci entraînerait l'impossibilité de sélectionner le programme moteur correct [bleu]. L'action du module de contrôle exécutif consiste, dans le cas d'un stimulus compatible, à augmenter le niveau d'activité du concept sélectionné par la couleur du mot jusqu'à ce que le lemme associé à ce concept atteigne le seuil de déclenchement de la réponse.

Le modèle prédit donc que le CCA est plus actif en condition incompatible qu'en condition compatible. La différence d'implication de ce module du contrôle exécutif entre les conditions compatible et incompatible se résume à la présence ou non de la suppression du traitement de l'attribut non-pertinent.

### Données expliquées par le modèle

Roelofs (2003) a montré que WEAVER ++ est capable de rendre compte des données comportementales obtenues dans la tâche de Stroop (pour une revue, voir MacLeod, 1991), notamment l'effet de compatibilité sur le temps de réaction ainsi que l'absence d'effet Stroop inverse (interférence de la couleur du mot sur la lecture du mot).

Lorsque la tâche est d'identifier la couleur, le temps de réponse est plus rapide en condition compatible puisque le lemme correct a un niveau d'activation qui a été augmenté par l'information en provenance du mot, ce qui diminue le nombre de cycles nécessaire à l'activation lemme correct pour franchir le seuil de réponse. L'absence d'effet Stroop inverse est expliquée par le fait que le lemme sélectionné par la lecture du mot franchi plus rapidement le seuil de déclenchement de la réponse que le lemme lié au concept activé par la couleur du mot.

Roelofs et Hagoort (2002) ont cherché à modéliser la réponse BOLD correspondant à l'activité du CCA dans le réseau et à comparer cette réponse BOLD-simulée aux données expérimentales issue d'une étude de Carter *et al.* (1995) et d'une étude MacDonald *et al.* (2000).

À partir des paramètres existants dans le modèle, Roelofs et Hagoort (2002) ont fait l'hypothèse que la réponse BOLD pouvait être décrite de cette façon :

$$\text{BOLD response } (t) = w t^{du/\Delta t} e^{-t/critdiff}, \quad (2.6)$$

où  $w$  correspond à la somme des entrées reçues par le module du contrôle exécutif, il est un indice

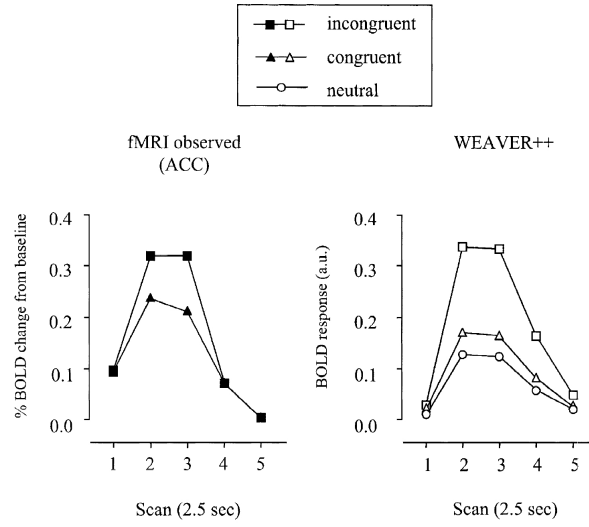


FIGURE 2.13 – Comparaison entre les données expérimentales issues de l’étude de MacDonald *et al.* (2000) et la simulation de la réponse BOLD dans le modèle de Roelofs (Roelofs et Hagoort, 2002)

de la quantité de contrôle nécessaire en fonction de la condition dans laquelle le sujet se trouve. Les paramètres  $du$  et  $critdiff$  correspondent respectivement à la latence avec laquelle le contrôle se met en place et au seuil de réponse dans le modèle. De façon simplifiée, les paramètres  $du$  et  $critdiff$  sont des constantes qui affectent la durée de la réponse BOLD simulée, tandis que le paramètre  $w$  affecte son amplitude.

“BOLD response” est une fonction dont la dérivée s’annule en  $t = critdiff$ . Elle est croissante sur l’intervalle  $[0, critdiff]$  et exponentiellement décroissante entre  $[0, +\infty[$ . Sur la figure 2.13 de gauche, on peut voir la réponse hémodynamique au niveau du CCA dans une tâche de Stroop issue de l’étude de MacDonald *et al.* (2000). Les résultats montrent que l’amplitude de la réponse hémodynamique est plus importante dans la condition incompatible que dans la condition compatible. La réponse BOLD simulée par le modèle WEAVER ++ selon l’équation 2.6 (figure 2.13 de droite), présente des résultats qui vont dans le même sens que ceux obtenus par MacDonald *et al.* (2000). En ce qui concerne l’amplitude de la réponse BOLD dans la condition neutre, la simulation du modèle montre que la réponse BOLD est plus faible en condition neutre qu’en condition compatible. Ce résultat est en accord avec les résultats obtenus par Carter *et al.* (1995) en tomographie par émission de positons. Le modèle WEAVER ++ est également capable de rendre compte de l’effet obtenu sur les temps de réaction observé par Carter *et al.* (1995). Les sujets sont les plus rapides en condition compatible et les plus lents en condition incompatible. Les TR des essais neutres sont

intermédiaires.

Dans une adaptation du modèle de WEAVER ++ à une tâche de Stroop modifiée combinant un mot et une flèche, Roelofs *et al.* (2006) ont montré que le modèle est capable de rendre compte des données acquises en IRMf dans cette tâche. Les auteurs ont comparé l'activité du CCA dans une tâche de compatibilité spatiale dans la condition compatible, incompatible et neutre (voir figure 2.14). Les résultats expérimentaux et ceux issus de la simulation montrent que le CCA a une activité importante dans la condition incompatible, intermédiaire dans la condition neutre et plus faible dans la condition compatible. Le modèle du conflit (Botvinick *et al.*, 2001) prédit que l'activité du CCA devrait être équivalente dans la condition compatible et dans la condition neutre. Les résultats de l'étude de Roelofs *et al.* (2006) sont donc incompatibles avec les prédictions du modèle du conflit. D'après les auteurs, le CCA pourrait avoir un rôle de régulateur.

Nous remarquons que les prédictions du modèle de WEAVER ++ sont différentes dans la tâche de Stroop classique et dans la tâche de Stroop spatiale (figures 2.11 et 2.14). Dans la première simulation, le contrôleur exécutif serait plus actif en condition compatible qu'en condition neutre, alors que dans la deuxième simulation le contrôleur exécutif serait plus actif en condition neutre qu'en condition compatible. Seuls les résultats de la première simulation sont cohérents avec ceux obtenus par Carter *et al.* (1995).

#### **Bilan du modèle Weaver ++ :**

Ce modèle est, comme le modèle du conflit, conçu avec deux modules indépendants. D'une part, un module chargé d'implémenter la tâche et d'autre part un module de contrôle. Celui-ci intervient en biaisant les informations en provenance du stimulus. Les informations non pertinentes sont bloquées tandis que les informations pertinentes ont leur activation renforcée. Cependant la mise en place de ce contrôle semble être automatique à chaque essai, et aucun ajustement entre les essais n'est envisagé. Ce modèle ne semble donc pas être capable de rendre compte des effets séquentiels.

Ce modèle est capable d'expliquer les données IRMf. Avec les paramètres du modèle, une équation permet de simuler cette réponse hémodynamique. Cependant, les auteurs de WEAVER ++ n'ont pas proposé d'équation permettant de rendre compte de l'activité électroencéphalographique.

Les auteurs ont montré que le modèle initial de WEAVER ++ peut être transformé et adapté afin qu'il puisse rendre compte de tâches manipulant la compatibilité stimulus-réponse qui appartiennent à la même classe de tâches selon Kornblum *et al.* (1990). En effet, WEAVER ++ ne prévoit pas d'étape d'association Stimulus-Réponse. Dans les deux tâches, la cible et les distracteurs n'en-

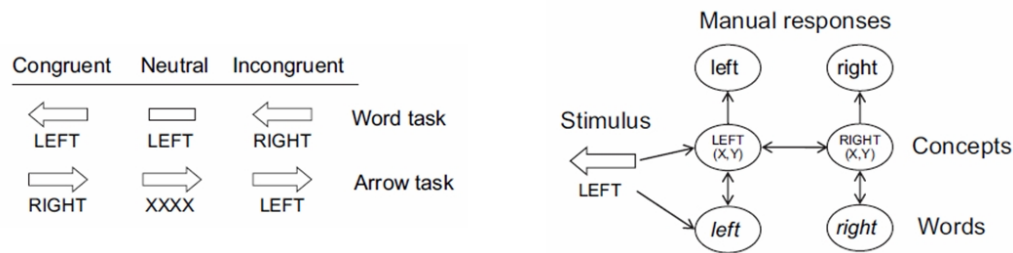


FIGURE 2.14 – **Illustration de la tâche et du modèle utilisés par Roelofs *et al.* (2006).**  
Figures tirées de Roelofs *et al.* (2006)

treignent pas de relation arbitraire avec les réponses.

Pour que *WEAVER ++* puisse rendre compte des données comportementales, son architecture doit tenir compte de la vitesse de traitement des deux attributs du stimulus. Les unités chargées de traiter l'attribut pertinent doivent être plus proches dans le réseau de la réponse que les unités de traitement chargées du traitement des distracteurs. La modélisation de la tâche de Stroop tient compte du fait que la lecture de mot est plus rapide que l'identification des couleurs (figure 2.14). En effet, la distance à parcourir dans le réseau est plus importante pour aller du traitement de la couleur à la production vocale que de la lecture du mot à la production vocale. De la même façon, la modélisation de la version spatiale de la tâche de Stroop tient compte du fait la lecture du mot est moins rapide que l'identification de l'orientation de la flèche (Roelofs *et al.*, 2006).

Le modèle *WEAVER ++* présente des avantages et des inconvénients. D'une part, il semble être adaptable aux tâches de type Stroop manipulant la compatibilité Stimulus-Réponse en prenant en compte la rapidité de traitement des différents attributs des stimuli. En revanche, il est nécessaire qu'il soit modifié pour qu'il puisse intégrer des règles d'association Stimulus-Réponse afin de pouvoir modéliser les tâches de type Simon ou Eriksen qui correspondent à des taxonomies différentes selon Kornblum *et al.* (1990). D'autre part, *WEAVER ++* est capable d'expliquer la réponse BOLD enregistrée au niveau du CCA, mais pas l'onde d'erreur. Enfin, il ne semble pas pouvoir rendre compte des ajustements comportementaux. Cependant, une adaptation du modèle pour qu'il puisse en rendre compte semble possible en intégrant ce qui s'est passé à l'essai précédent. On pourrait imaginer que le contrôleur exécutif garde une trace de son activité à l'essai précédent.

## 2.3 Modèles Ganglions de la Base

Nous avons vu au paragraphe 1.2.2 page 20 que les ganglions de la base sont sollicités lors des apprentissages par renforcement. Si l'on considère que les activités électroencéphalographiques FRN et  $N_E$  sont similaires et observées respectivement en phase d'apprentissage et en phase de test des apprentissages par renforcement, et si la FRN est bien le reflet des variations d'activité du système mésencéphalique dopaminergique, alors il est important d'intégrer les GDB au sein des processus de contrôle des actions. Cette idée a été utilisée par Holroyd et collaborateurs ainsi que par Frank et collaborateurs dans respectivement le modèle d'apprentissage par renforcement et le modèle des ganglions de la base que nous décrivons ci-dessous (Holroyd et Coles, 2002; Holroyd *et al.*, 2005; Frank, 2005).

### 2.3.1 Modèle d'Apprentissage par Renforcement

Le but d'Holroyd et collaborateurs était de lier la présence de la FRN au cours de l'apprentissage et de la  $N_E$  une fois l'apprentissage terminé à des changements d'activité du système mésencéphalique dopaminergique liés<sup>1</sup> à la prédiction de la récompense (sDA, voir paragraphe 1.2.2 page 20). D'autre part, ils cherchaient à développer un modèle neuromimétique de l'apprentissage par renforcement des associations Stimulus-Réponse dans des tâches de temps de réaction de choix (Holroyd et Coles, 2002; Holroyd *et al.*, 2005). L'activité du système mésencéphalique dopaminergique (sDA) a été décrite au paragraphe 1.2.2 page 20.

L'idée générale de la théorie de l'apprentissage par renforcement ("Reinforcement Learning, theorie" - ERN-RL<sup>2</sup>) est que la FRN et la  $N_E$  sont le reflet de la transmission du signal d'erreur émis par le sDA au CCA. En début d'apprentissage, seul le feedback permet de savoir si la réponse était correcte ou erronée. C'est donc le feedback qui véhicule le signal d'erreur. Lorsque qu'un feedback négatif est donné aux sujets, on enregistre une FRN environ 200 ms après la présentation du feedback. Par contre, dans la phase de test, au moment où le sujet produit une réponse erronée, il sait que c'est une erreur. C'est donc la réponse qui conduit au signal d'erreur et on observe une  $N_E$  environ 70 ms après la réponse. Pour Holroyd *et al.* (2005), la  $N_E$  et la FRN sont deux ondes qui reflètent le même processus : la prédiction d'une erreur par le système d'évaluation des comportements.

Comme pour le modèle du conflit et WEAVER ++, le modèle ERN-RL se caractérise par deux

---

1. Ces deux activités électroencéphalographiques  $N_E$  et FRN ont été décrites aux paragraphes 1.3.1 et 1.3.2 pages 23 et 24.

2. Nous avons conservé l'abréviation originale du modèle. ERN et  $N_E$  se réfèrent à la même activité

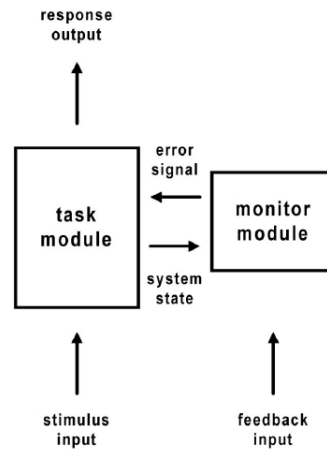


FIGURE 2.15 – **Modèle ERN-Reinforcement Learning simplifié proposé par Holroyd et al. (2005)**

modules bien distincts qui fonctionnent en parallèle (voir figure 2.15). L'un des modules gère la tâche, tandis que l'autre est chargé de l'évaluation et de la surveillance des actions effectuées. Les originalités de ce modèle consistent dans la conception du module de d'évaluation, dans la signification du signal d'erreur ainsi que dans son utilisation.

#### Les grandes hypothèses :

1. Le module d'évaluation agit comme un comparateur entre la récompense espérée et la récompense reçue. Cette comparaison émet un signal d'erreur qui peut être positif lorsque la récompense espérée est meilleure que la récompense reçue, négatif lorsque la récompense espérée est moins bonne que la récompense reçue, ou nul lorsque l'estimation de la récompense est bonne.
2. Le signal d'erreur est utilisé pour optimiser la performance à la tâche ainsi que pour améliorer la prédiction des performances.

Les données expérimentales et les résultats des simulations ont conduit les auteurs à réinterpréter l'activité de la  $N_E$  et de la FRN observées en EEG comme étant la manifestation neurophysiologique d'un signal d'erreur négatif véhiculé par le sDA.

#### Fontionnement général du modèle

Pendant la phase d'apprentissage, le module tâche identifie le stimulus et donne une réponse en fonction de la nature de ce stimulus. En parallèle, le module d'évaluation récupère les infor-



mations sur la réponse sélectionnée et le programme moteur exécuté, et “propose” une évaluation du comportement choisi. Il s’agit d’une valeur entre -1 et 1 correspondant au “risque” d’obtenir la récompense<sup>3</sup>. À la fin de l’essai, un feedback positif ou négatif est délivré.

Le module évaluation mène alors deux opérations :

- Le feedback permet de savoir si l’évaluation du comportement exécuté par le module tâche est correcte.
- Le feedback permet de savoir si l’appariement stimulus-réponse choisi par le module tâche est correct.

Plus précisément, le module d’évaluation compare sa propre prédiction au feedback reçu. Cette comparaison génère un signal d’erreur. Trois cas de figure sont possibles :

1. le signal d’erreur est supérieur à 0 : la bonne réponse a été donnée et la récompense est meilleure que l’estimation. Les paramètres du module évaluation vont être modifiés alors que ceux du module tâche seront conservés.
2. le signal d’erreur est inférieur à 0 : La réponse donnée est incorrecte et l’évaluation du comportement également. Les paramètres du module évaluation et du module tâche devront être modifiés.
3. le signal d’erreur est égal à 0 : l’estimation de la récompense est bonne. Si la valeur est proche de 0, on peut savoir si la réponse délivrée est correcte ou incorrecte, et modifier le module tâche en cas de réponse incorrecte.

Cet algorithme d’apprentissage permet au module tâche d’améliorer ses performances et au module d’évaluation améliorer l’estimation des comportements du module tâche. En phase de test, le feedback n’est plus délivré. Le module évaluation a appris à reconnaître les comportements corrects et peut ainsi évaluer les comportements délivrés par le module tâche.

### Description du module tâche

Le module tâche est composé de plusieurs couches (figure 2.16 de gauche). Ce module ressemble beaucoup au module tâche utilisé par le modèle du conflit (voir figure 2.6 page 51). La couche des stimuli est chargée de coder les différents attributs des stimuli. Elle est composée de 6 unités : deux unités sont nécessaires pour coder le distracteur de gauche, deux autres pour coder le distracteur de droite, et encore deux autres pour coder la cible. Chacune de ces unités est reliée aux unités

---

3. La valeur absolue ce “risque” correspond à la probabilité et le signe à la valence. Une valeur de -1 correspond à la certitude de ne pas obtenir la récompense et la valeur de 1 correspond à la certitude de l’obtenir

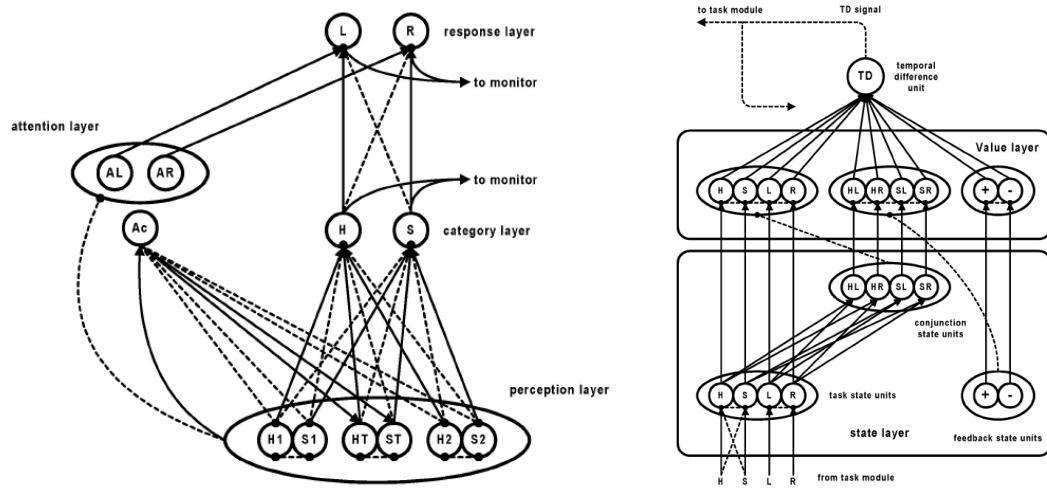


FIGURE 2.16 – Modèle ERN-Reinforcement Learning proposé par Holroyd *et al.* (2005) adapté pour la tâche d'Eriksen. Sur l'illustration de gauche le module tâche et sur celle de droite le module évaluation

de la couche catégorielle (*category layer*). Cette couche est chargée de sélectionner la solution en fonction des informations en provenance de la couche des stimuli. Lorsque l'activation de l'une des unités dépasse un seuil, alors l'unité catégorielle dont la solution a été sélectionnée active l'une des unités de la couche de réponse. La couche de réponse est constituée de plusieurs unités de réponse correspondant aux différents effecteurs moteurs (par exemple appui sur un bouton avec le pouce droit ou le pouce gauche). L'association stimulus-réponse est codée grâce aux poids des connexions entre la couche catégorielle et la couche de réponse. Ces trois niveaux de couches correspondent à trois étapes dans les processus de traitement de l'information et d'organisation fonctionnelle des neurones dans le cerveau : des populations de neurones sont spécialisées dans l'encodage des stimuli, d'autres dans l'identification de la cible et d'autres dans la production de la réponse. Les règles d'association S-R sont implémentées dans les connexions entre l'étape d'identification de la cible et l'étape de production de la réponse.

Enfin la couche attentionnelle facilite les processus d'identification de la cible en augmentant le niveau d'activité des unités chargées du traitement des caractéristiques pertinentes du stimulus pour la tâche et en abaissant le niveau d'activité des unités chargées du traitement des caractéristiques non-pertinentes du stimulus pour la tâche. Le module attentionnel implémente également des biais de réponse au niveau de la couche de réponse.

### Description du module évaluation

Le module d'évaluation est composé de trois couches (figure 2.16 à droite).

Une première couche appelée couche "état" est constituée de plusieurs groupes de neurones. Les informations sur la solution sélectionnée par la couche catégorielle et sur la réponse exécutée dans le module tâche, sont recueillies par un groupe de neurones spécifique. Dans la phase d'apprentissage, l'information sur le feedback est utilisée dans cette couche du réseau. L'originalité de ce modèle tient dans l'existence d'une couche particulière appelée "couche des unités de conjonction" qui détectent l'association S-R utilisée par le module tâche pour répondre.

La couche des "valeurs" récupère les informations sur la cible identifiée, la réponse choisie, la valeur de la prédiction (entre -1 et 1) correspondant à l'association S-R utilisée pour répondre et le feedback.

Enfin, l'unité de différence temporelle compare la récompense espérée avec le feedback et calcule le signal d'erreur qui sera utilisé par le module tâche et le module d'évaluation. Ce signal d'erreur peut être positif ou négatif en fonction du sens de l'erreur de prédiction : le signal d'erreur est positif (ou négatif) lorsque la récompense reçue est meilleure (ou pire) que celle espérée.

**Dynamique du Modèle** Au début de la phase d'apprentissage, les comportements du module tâche et du module évaluation n'obtiennent pas de bons résultats. Au fur et à mesure de la phase d'apprentissage, les prédictions des unités de conjonction vont s'améliorer et le module tâche va moins commettre d'erreurs. Les informations sur le feedback vont être de moins en moins utilisées. Holroyd *et al.* (2005) prédisent donc une diminution au fur et à mesure de la tâche de l'amplitude de la FRN et une augmentation de l'amplitude de la  $N_E$ .

### Bilan sur le modèle d'apprentissage par renforcement

Ce modèle est intéressant de plusieurs points de vues. D'une part, il fait un lien explicite entre la  $N_E$  et la FRN. Par cette occasion, il inclut la dimension de l'apprentissage. Nous avons vu qu'il y a des liens importants entre la mise en place du contrôle cognitif au cours du développement. Ce modèle permet de faire le lien entre différents champs de la littérature développement, GDB, Apprentissage, FRN/Ne.

Puisqu'il suppose que le système dopaminergique mésencéphalique est impliqué dans ces processus, ce modèle devient également testable avec une approche pharmacologique qui pourrait être menée chez l'Homme et également chez l'animal.

Comme les autres modèles, il ne peut pas rendre compte de la  $N_E$  dans les essais corrects.

### 2.3.2 Modèle de Ganglions de la Base

Frank (2005) a construit un modèle neuromimétique chargé de mimer le fonctionnement des ganglions de la base (GDB) ainsi que les connexions avec les cortex prémoteur/préfrontal et le thalamus dans des tâches cognitives impliquant l'attribution de récompenses.

La construction du modèle neuromimétique est basée sur le modèle de fonctionnement des GDB suivant (illustré sur la partie gauche de la figure 2.17). Les neurones dopaminergiques de la substance noire compacte (SNc) projettent sur les neurones striataux. Il existe deux sous populations de neurones striataux qui diffèrent par leur cible et leur sensibilité à la dopamine :

- **La voie dite “directe”** concerne les neurones striataux GABAergiques, porteurs de récepteurs dopaminergiques D1, et qui projettent sur le globus pallidus interne (GPi) et la substance noire réticulée (SNr).
- **La voie dite “indirecte”** concerne neurones striataux GABAergiques, porteurs de récepteurs dopaminergiques D2, et qui projettent sur le GPi et la SNr via un relais : le globus pallidus externe (GPe).

Ces deux voies ont des effets opposés sur la sortie des ganglions de la base (DeLong, 1990) et donc des rôles fonctionnels différents. L'activation de la voie directe induit une diminution de l'activité des GPi/SNr, conduisant à une désinhibition de l'activité thalamocorticale dont les conséquences sont une facilitation des mouvements. Les neurones du striatum de la voie directe sont nommés les neurones “Go” par Frank (2005). À l'inverse, l'activation de la voie indirecte aboutit à une hyperactivité des GPi/SNr, donc à une inhibition de l'activité thalamocorticale ce qui tend à réduire les mouvements. Les neurones du striatum de la voie indirecte sont nommés les neurones “No-Go” par Frank (2005).

L'illustration de droite de la figure 2.17 représente la modélisation neuromimétique des GDB de Frank (2005). Ce modèle est composé de plusieurs couches, chacune d'elles représente une structure nerveuse correspondant au modèle simplifié de fonctionnement des GDB. Ce modèle fonctionne de la façon suivante. L'arrivée d'un stimulus correspond à une configuration particulière des unités dans la couche des entrées (*input*). Cette configuration stimule, par un système de poids entre les connexions, les unités (ou neurones) Go ou No-Go du striatum, qui, à leur tour vont activer la voie directe ou indirecte des GDB. Le thalamus est soit activé, soit inhibé. En cas d'activation du thalamus par la voie directe, le cortex prémoteur activé va permettre le mouvement. En cas d'inhibition du thalamus par la voie indirecte, le mouvement ne sera pas exécuté.

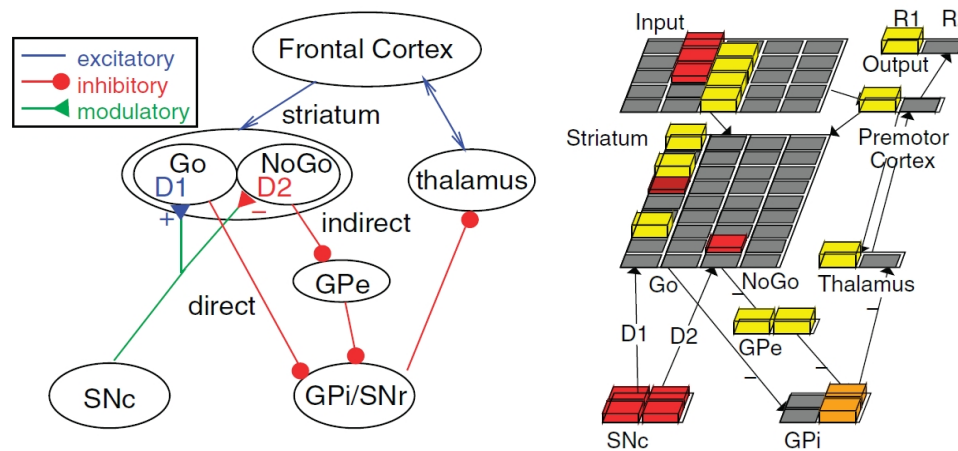


FIGURE 2.17 – **Modélisation du fonctionnement des ganglions de la base utilisée par Frank (2005) pour la construction du modèle neuromimétique des GDB.** Figures tirées de Frank *et al.* (2004). (SNc : substance noire compacte ; SNr : substance noire réticulée ; GPe : globus pallidus externe ; GPi : globus pallidus interne). Les traits bleus, rouges et verts symbolisent respectivement les excitations, les inhibitions et les modulations.

### Modèle des GDB et apprentissage

Dans la littérature animale, il a été montré que la dopamine jouait un rôle majeur dans les apprentissages par renforcement (voir paragraphe 1.2.2 page 20). Des bouffées phasiques de DA ont été enregistrées pendant des renforcements positifs et des diminutions phasiques de DA lors de renforcements négatifs (Schultz *et al.*, 1997; Schultz, 2002; Satoh *et al.*, 2003). L'apprentissage a été modélisé de la façon suivante dans le modèle des GDB. À la fin de l'essai, un feedback positif ou négatif est véhiculé par la SNc.

- **Si le feedback est positif**, la SNc augmente la libération de DA vers le striatum. La conséquence est une augmentation des poids synaptiques entre les unités de la couche d'entrée et les neurones Go, et une diminution des poids synaptiques entre les unités de la couche d'entrée et les neurones No-Go. Le réseau "apprend" Go lors d'un renforcement positif, ce qui facilite le mouvement.
- **Si le feedback est négatif**, la SNc diminue la libération de DA vers le striatum, ce qui entraîne une diminution des poids synaptiques entre les unités de la couche des entrées et les neurones Go et une augmentation des poids synaptiques entre les unités de la couche des entrées et les neurones No-Go. Le réseau "apprend" No-Go pour éviter les choix erronés.

Lorsque l'apprentissage est terminé, les poids synaptiques entre la couche des entrées et les neurones du striatum sont fixés. Chaque stimulus est associé plutôt la voie directe GO ou la voie indirecte

No-Go.

### Prédictions sur l'influence de la DA sur l'apprentissage au sein des GDB

Frank *et al.* (2004) ont mimé un déficit en DA dans le modèle. Ils ont constaté des prédictions du modèle non triviales. La capacité d'apprentissage du modèle n'est pas nulle, mais seulement diminuée. Le réseau parvient en effet à utiliser les informations sur les feedback négatifs, mais pas sur les feedbacks positifs. Plus précisément, le niveau général de DA étant abaissé, l'augmentation phasique du niveau de DA due à l'arrivée d'un feedback positif n'est plus suffisante pour être détectée par les neurones du striatum (effet de seuil). En revanche, la diminution phasique déclenchée par une feedback négatif du niveau de DA reste détectable. Si ces prédictions sont exactes alors l'expression "marcher à la carotte ou au bâton"<sup>4</sup> prend tout son sens.

Si cette prédiction est vraie, alors les patients souffrants de déficits en DA devraient réussir à apprendre une tâche en utilisant les informations liées aux feedback négatifs. Les patients atteints de la maladie de Parkinson (patients MP) présentent une dégénérescence de la substance noire et par conséquent leur niveau de DA est diminué. Des déficits cognitifs, notamment dans les apprentissages par renforcement, sont visibles. Frank *et al.* (2004) ont testé l'hypothèse selon laquelle les patients MP ne devraient pas être aidés par les informations véhiculées par les feedbacks positifs, mais celles véhiculées par les feedbacks négatifs devraient leur permettre un apprentissage. En revanche, les patients MP prenant un traitement substitutif au manque de DA devraient réussir à apprendre de leurs feedback positifs puisque le niveau de DA est rehaussé. La hausse de la DA déclenchée par un feedback positif, serait facilement détectable par les neurones du striatum. Par contre, la médication empêcherait au niveau de DA de baisser suffisamment pour que le feedback négatif soit perçu au niveau du striatum<sup>5</sup>. Frank *et al.* (2004) a conçu une tâche permettant tester explicitement cette hypothèse.

### Interprétation de la Ne dans ce modèle

Pour Frank *et al.* (2005) comme pour Holroyd *et al.* (2005), la  $N_E$  est le témoin de la diminution de la fréquence de décharge des neurones DA de la SNc lorsque le sujet reçoit un feedback négatif.

Dans le modèle des GDB, la baisse de DA témoigne d'un apprentissage No-Go, les sujets apprennent à éviter cette réponse. Après avoir trié les sujets en fonction du type de feedback

---

4. Cette expression est la traduction française de l'expression anglaise "*By carrot or by stick*" utilisée comme titre du papier de Frank *et al.* (2004).

5. Il y aurait un effet de type plancher.

auquel ils sont sensibles, les auteurs ont montré que les sujets utilisant davantage les informations sur renforcements positifs pour apprendre la tâche (sujets r+) présentaient une  $N_E$  plus petite que les sujets utilisant davantage les informations sur les renforcements négatifs (sujets r-). Ils ont interprété ces résultats de la façon suivante : l'amplitude de la  $N_E$  permet de prédire le degré de prise en compte des feedback négatifs par le sujet. Les sujets r- ont une  $N_E$  plus grande, ce qui signifie qu'ils sont davantage sensibles aux erreurs et apprennent mieux grâce à elles. Ils ont regardé la FRN et ont noté qu'elle est plus grande au moment des feedback négatifs qu'au moment des feedback positifs chez les sujets r- uniquement. Les résultats vont donc dans le sens que la  $N_E$  et la FRN reflètent des processus neuraux communs au niveau du CCA associés à une baisse de libération de DA.

### Bilan du modèle des GDB

L'idée que certaines personnes "marchent à la carotte et d'autres au bâton" est familière et son utilisation pour expliquer scientifiquement les comportements est assez séduisante. La force du modèle des GDB est de s'être montré capable de prédire un effet du niveau de DA sur l'apprentissage qui a été observé ensuite dans une étude expérimentale (Frank *et al.*, 2004). L'idée de chercher des différences physiologiques sur la base des performances des sujets a permis de montrer un effet de l'amplitude de la  $N_E$ . Les sujets apprenant davantage des feedbacks négatifs présentent une  $N_E$  de plus grande amplitude que les sujets apprenant davantage des feedbacks positifs.

Cependant, une étude très récente déstabilise l'une des hypothèses fondamentales du modèle des GDB. Willemssen *et al.* (2009) ont comparé l'amplitude de la  $N_E$  chez les patients *de novo* atteints de la maladie de Parkinson. La maladie venant d'être dépistée chez ces patients, aucun traitement compensant la baisse du niveau de DA n'a été administré. L'amplitude de la  $N_E$  dans les erreurs est plus faible chez ces patients que chez les patients contrôle du même âge. Or l'hypothèse de Frank et collaborateurs prédit que le niveau de base de DA conditionne le mode d'apprentissage des sujets et que ce mode d'apprentissage peut être détectée en regardant l'amplitude de la  $N_E$  qui est révélatrice de la stratégie des sujets. Le niveau de DA étant plus faible chez les patients *de novo* que chez les sujets contrôles, ils devraient présenter une  $N_E$  de plus grande amplitude comparée à celle des sujets contrôles. Toutefois, la comparaison de ces deux études est délicate. Les deux tâches utilisées sont de nature différente. Willemssen *et al.* (2009) ont utilisé une tâche d'Eriksen modifiée sans recherche de règle tandis que la tâche utilisée par Frank et collaborateurs est une tâche dans laquelle le sujet doit découvrir la bonne réponse. Les résultats obtenus par Frank *et al.* (2004) chez les parkinsoniens sont des résultats comportementaux. L'idée selon laquelle la  $N_E$  devrait être

plus grande chez les patients que chez les contrôles est une extrapolation des données de l'étude de Frank *et al.* (2005) chez les sujets sains qui rapporte que la  $N_E$  est plus ample chez les sujets sensibles aux informations sur les feedbacks négatifs. L'idéal pour tester correctement le modèle des GDB serait de tester des patients novo dans la tâche imaginée par Frank *et al.* (2004), et de s'intéresser à la fois aux données comportementales et électroencéphalographiques.

## 2.4 Conclusion sur les modèles

La grande diversité des modèles met en lumière la complexité des processus de contrôle. Tous ces modèles ont des avantages et des inconvénients. Leur point commun est de chercher à rendre compte de l'activité du CCA dans les tâches qui nécessitent un contrôle cognitif. Les modèles du conflit, ERN-RL et des GDB proposent des interprétations différentes de la  $N_E$ . Tandis que WEAVER ++ offre une modélisation de la réponse hémodynamique du CCA enregistrée en IRMf. Le tableau 2.18 présente les différences observées entre les différents modèles.

Le modèle du conflit rend compte à la fois de la  $N_E$  dans les erreurs et de la N2 dans les essais corrects (Botvinick *et al.*, 2001). Il a l'énorme avantage de proposer un moyen économique pour détecter les situations dans lesquelles le contrôle cognitif doit entrer en jeu. En se "contentant" de regarder si plusieurs activités sont activées en même temps, il contourne la principale difficulté rencontrée par les modèles, qui est de savoir quelle est la réponse correcte (ces modèle font la comparaison entre la réponse correcte et la réponse en cours d'exécution). Cette capacité rend le modèle du conflit adaptable à toutes les tâches dans lesquelles il faut faire un choix. Bien qu'il n'ait pas proposé de modélisation de la réponse hémodynamique, ce modèle est également cohérent avec les nombreuses études IRMf montrant des augmentations d'activité dans le CCA et le DLPFC après les essais dans lesquels il y a eu un risque d'erreur (Kerns *et al.*, 2004). Par contre, il ne peut ni expliquer la FRN, ni expliquer la  $N_E$  dans les essais corrects.

Le modèle ERN-RL rend compte de la FRN et la  $N_E$  (Holroyd *et al.*, 2005). Il fait le lien entre les processus d'apprentissages et les connexions entre les ganglions de la base et le CCA. La FRN et la  $N_E$  sont proposées comme étant le fruit d'un processus d'évaluation de la réponse basé sur l'évaluation de l'association S-R utilisée pour donner la réponse. Les erreurs détectées sont des erreurs d'association S-R. En revanche, il ne peut pas rendre compte de la  $N_E$  des essais corrects et de la N2. Ce modèle est beaucoup plus coûteux que le modèle du conflit. En effet, pour évaluer la réponse, le modèle prévoit autant d'unités de conjonction nécessaires que d'associations stimulus-réponse possibles. Si  $x$  correspond au nombre de réponses possibles et  $y$  le nombre de



cibles possibles, le module évaluation a besoin de  $x \times y$  unités de conjonctions. Cet aspect pose un problème notamment dans les erreurs de langage pour lesquelles le nombre de réponses possibles est bien trop important pour que toutes les associations S-R soient représentées par des unités de conjonction.

Le modèle *WEAVER ++* proposé par Roelofs (2003) est capable d'expliquer des ajustements comportementaux en ligne au moyen d'un contrôleur exécutif détectant un risque d'erreur. Il propose que le CCA entre en jeu en cours d'essai et non à la fin de l'essai comme dans les modèles du conflit et ERN-RL. Il aurait un rôle de régulation sur le traitement perceptifs des attributs du stimulus en cours d'essai. *WEAVER ++* est capable de modéliser la fonction hémodynamique, mais pas les activités EEG. L'activité du CCA précédant la réponse dans les essais incompatibles pourrait correspondre à l'activité N2 observée en EEG calé sur la réponse. Toutefois, la signification fonctionnelle donnée à cette N2 ne correspond pas aux hypothèses faites par *WEAVER ++*.

Le modèle des GDB proposé par Frank (2005) est à part des autres modèles pour plusieurs raisons. Son système de contrôle est intrinsèquement lié à la prise de décision alors que pour les autres l'évaluation est indépendante du choix de la réponse. Puisque que la sortie de ce modèle est Go ou No-Go, il semblerait qu'un choix de réponse ne soit pas faisable directement au sein des GDB. Pour autant, nous avons l'impression de ce modèle pourrait parfaitement correspondre à un module d'évaluation qui s'adapterait sur une modélisation d'une tâche. Dans ce cas, il ne serait pas très loin conceptuellement du module évaluation du modèle ERN-RL.

Aucun de ces modèles ne peut rendre compte de la  $N_E$  dans les essais corrects. Cette activité pose un problème pour la modélisation. En effet, si la  $N_E$  des corrects reflète le même processus physiologique que la  $N_E$  des erreurs et la FRN, alors nous devons tenter de trouver une explication pour rendre compte de ces trois activités.

	Évaluation	S-R	CCA	$N_E$	FRN	N2	$N_E$ des corrects
Modèle du conflit	séparé			×		×	
WEAVER ++	séparé		×				
ERN-RL	séparé	×		×	×		
GDB	inclu	×		×	×		

FIGURE 2.18 – **Tableau représentant les différents aspects des modèles neuromimétiques.**  
 Évaluation : le modèle prévoit-il une évaluation en dehors de son système de sélection de la réponse ?  
 S-R : existe t-il une étape d'association Stimulus-Réponse ? CCA : les modèles peuvent-ils modéliser  
 l'activité IRMf enregistrée dans le CCA ?  $N_E$  , FRN, N2  $N_E$  des corrects : les modèles peuvent-ils  
 rendre compte de ces activités EEG ?

Deuxième partie

Partie Expérimentale



## Chapitre 3

# Error negativity does not reflect Conflict. A re-appraisal of conflict monitoring and Anterior Cingulate Cortex activity<sup>1</sup>

### 3.1 Introduction

In order to adapt to ever-changing environments, animals must continuously alter their behavior. Such flexibility is often assumed to be mediated by “control” mechanisms that adjust information processing to the prevailing context. The way in which control mechanisms are recruited, however, remains obscure. In the last few years, the “conflict-loop theory” (Botvinick *et al.*, 2001, 2004, 1999; Carter *et al.*, 1999, 1998; Cohen *et al.*, 2000; Yeung *et al.*, 2004) has played an essential role in this field by providing a unified model that aims to account for both neurophysiological and behavioral aspects of control implementation. This model introduced a very simple, though very powerful, concept, namely “response conflict”. Response conflict, measured by the Anterior Cingulate Cortex (ACC), is explicitly defined as the product of the activation of the responses weighted by the inhibitory connections between these responses (Botvinick *et al.*, 2001; Yeung *et al.*, 2004, see also eq. 3.2). In this type of neural network model, “response activation”

---

1. article publié dans la revue *Journal of Cognitive Neuroscience* 20 :9, pp. 1637–1655

refers to the amount of neural activity in the structures involved in response execution. Conflict monitoring has been largely studied in the so-called flanker task (Eriksen et Eriksen, 1974), in which participants must issue a right or a left hand response as a function of a target letter (*e.g.* the letters S or H), flanked by distractors that can be compatible (SSS) or incompatible (HSH) with the target. The implementation of the conflict model for the flanker task has been extensively described by Botvinick *et al.* (2001) and Yeung *et al.* (2004). Only the major aspects, and those directly relevant for our purpose, will be described here.

The model is implemented as a three-layers neural network (Cohen *et al.*, 1992) : one perceptual layer codes for the target and both distractors, one response layer codes for the two competing responses, and an attentional layer biases processing towards the target. All between-layer connections are excitatory (no between-layer inhibition), whereas all within-layer connections are inhibitory. To this basic architecture, the conflict modelers added a conflict-monitoring unit, that measures online the amount of conflict. Generally speaking, the conflict is measured as the energy (Hopfield, 1982) in the response layer, defined as :

$$- \sum_i \sum_j a_i a_j w_{ij} \quad (3.1)$$

where  $a$  represents the activity of each unit in the layer, indices  $i$  and  $j$  index the different units, and  $w_{ij}$  is the inhibitory connection between the units  $i$  and  $j$ .

In the flanker task, where only two possible responses are present, conflict over time,  $Co(t)$ , is computed as :

$$Co(t) = \begin{cases} -2 \times (act_H(t) \times act_S(t) \times -3) & \text{if } act_H(t) > 0, act_S(t) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

where  $act_H(t)$  - respectively  $act_S(t)$  - is the activation level of the unit coding for the H response - respectively the S response - at time  $t$ , and -3 is the (constant) inhibitory connection between the left and the right response units.

Conflict monitoring is assumed to be measurable with fMRI (Barch *et al.*, 2000; Botvinick *et al.*, 1999; Carter *et al.*, 1998) and EEG (Gehring et Fencsik, 2001; Rodríguez-Fornells *et al.*, 2002; Van Veen et Carter, 2002; West, 2004; Yeung *et al.*, 2004). In the present study, we will concentrate on the EEG correlate : when participants commit an error, one observes a negative component, maximal fronto-centrally, called “Error Negativity” ( $N_e$ ) (Falkenstein *et al.*, 1991), or “Error Related Negativity” (ERN) (Gehring *et al.*, 1993), starting about 30 ms and peaking about 100 ms after the incorrect electromyographic (EMG) activity. Simulations of the model have revealed that conflict is maximal just after the incorrect response activation; its timing

is thus comparable to that of the  $N_E$ <sup>2</sup>. This similarity led Botvinick *et al.* (2001) and Yeung *et al.* (2004) to re-interpret the  $N_E$  in terms of conflict; conflict that results secondarily in error detection. According to these authors, the time course of the  $N_E$  should parallel the time course of the conflict. One way of testing this view is to estimate conflict on a trial-by-trial basis. Since conflict is defined as the degree of co-activation of the responses, it amounts to estimating the degree of overlap between correct and incorrect response activations. One major difficulty is that response activation is a covert process, not directly observable. However, EMG recordings of the muscles involved in responding have proved efficient in revealing such covert activation (*e.g.* Burle *et al.*, 2005, 2002b; Coles *et al.*, 1985; Hasbroucq *et al.*, 1999; Smid *et al.*, 1990) : on some trials, although the correct response was given, subthreshold EMG activity in the muscles involved in the incorrect response (“partial error”, see Figure 3.1A) can be observed. A  $N_E$  has been reported to occur just after such partial errors (Burle *et al.*, 2005; Masaki et Segalowitz, 2004; Vidal *et al.*, 2000). Since these partial errors are detectable on a trial-by-trial basis, they offer the unique opportunity to directly evaluate the activation of incorrect responses, not otherwise observable on overt performance.<sup>3</sup>

Importantly, conflict is defined as the degree of *co-activation* of the two responses (see eq. 3.2). We thus reasoned that it should depend on the temporal overlap between incorrect and correct response activations, that is on the time separating the incorrect from the correct response ( $\Delta$  value on fig. 3.1). In order to verify the validity of this reasoning, we first ran a simulation study in which trials containing the equivalent of partial errors were analyzed. Since, as the results will show, this assumption was supported, the  $\Delta$  value was the main variable of interest. We then compared the simulated data to real, experimental data (see below). In order to compare the simulation results with the experimental ones, we processed the simulation and the EEG data in the same way. To

---

2. According to the conflict model, conflict is maximal after the incorrect response on erroneous trials, but peaks just before the correct response on correct trials (see (Yeung *et al.*, 2004, Figure 14)). This is due to the fact that, on correct trials, the incorrect activation precedes the correct activation. It has been argued that, in this case, the electrophysiological correlates of the conflict is to be found on a “N2” component, either time-locked to the stimulus (Yeung *et al.*, 2004, Figure 15), or response-locked (Yeung *et al.*, 2004, Figure 13). As the simulations presented below will clearly show, this N2 in fact reflects a  $N_E$  time-locked to the incorrect response activation, as predicted by the conflict model.

3. Note that, although EMG-fMRI co-registration is now starting to become feasible, detecting such small incorrect EMG activations during fMRI acquisition is still beyond the current technical capabilities, since those partial errors are of small amplitude, and the frequency components of EMG largely overlaps the frequencies of the noise induced by the EPI. It is therefore currently technically impossible to measure the degree of response co-activation in the scanner. For this reason, EEG is more appropriate than fMRI to test this aspect of the model.

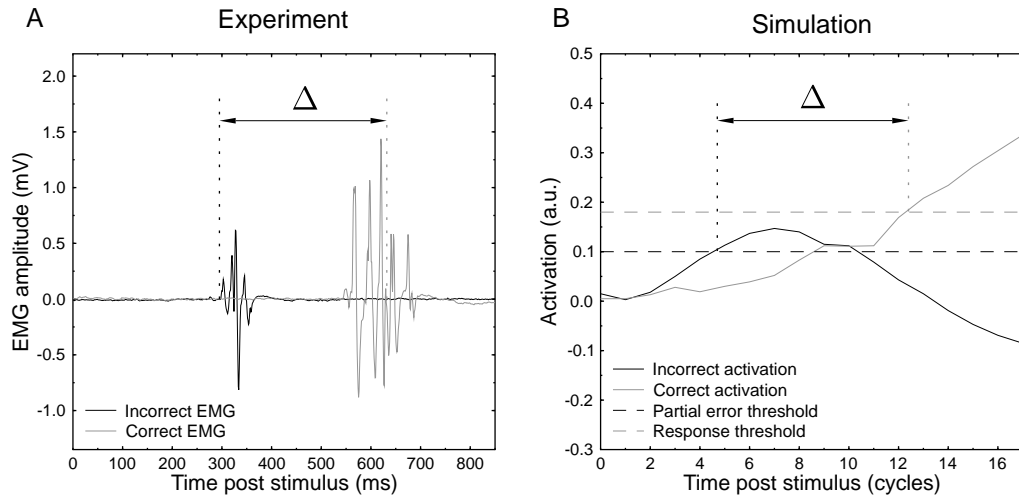


FIGURE 3.1 – Examples of “partial errors” obtained in the experiment and in the simulation. (A) The left part of the graph presents the electromyographic (EMG) activity as a function of time, in the muscles involved in the execution of the incorrect (black) and in the correct responses (grey). The time origin corresponds to stimulus presentation. The vertical black dashed line corresponds to the onset of the partial error, and the vertical grey dashed line to the mechanical response. Although the correct response was given, one can observe an activation of the incorrect response whose amplitude is, however, too low to trigger an overt error. The time interval between the incorrect EMG onset and the correct mechanical response is defined as  $\Delta$ , and will be used as an index of the temporal overlap between the two response activations. (B) The right part of the figure presents a simulated trial, with the amount of activation (in arbitrary units) of the incorrect (black line) and of the correct (grey line) response as a function of time. The time-zero corresponds to stimulus presentation, the vertical black dashed line corresponds to the onset of the partial error, and the vertical grey dashed line to the timing of the response. The two horizontal long dashed lines correspond to the partial error threshold (black) and to the correct response threshold (grey)



this end, we used the open-source software EEGLAB (Delorme et Makeig, 2004) which allows single-trial dynamics to be studied Jung *et al.* (2001) and hence reveals more precisely the impact of  $\Delta$  value on conflict and  $N_E$  amplitudes.

## 3.2 Methods

### 3.2.1 Simulation study

#### Simulation parameters

The simulation was based on 10 runs (representing 10 participants) of 1,000 trials each. All parameters used in the simulation were those used previously in the studies of Botvinick *et al.* (2001) and Yeung *et al.* (2004)<sup>4</sup>, except the threshold for “partial errors” (which were not previously considered). In the previous simulations, a response was recorded as soon as one of the two response units reached an activation level of .18. Using this value, a pilot simulation study indicated that a threshold of .10 for partial errors gives 10-15% partial errors, which is equivalent to the percentage usually obtained in empirical studies (Burle *et al.*, 2002b; Gratton *et al.*, 1988; Hasbroucq *et al.*, 1999; Smid *et al.*, 1990). Importantly, changing this value modified the overall number of partial errors, but did not affect the global performance pattern, as the difference in partial errors rate between compatible and incompatible trials remained constant whatever this value (though within realistic limits). Thus the results presented below are not specific to a well-tuned, somewhat arbitrary, value that we have chosen for partial errors threshold.

#### Data preprocessing

Depending on whether the correct or incorrect response unit reached the response threshold first, the trial was classified as correct or erroneous, respectively. Among the correct trials, we checked whether the incorrect response unit reached the partial error threshold before the correct response (it never happened that the incorrect response layer reached the partial error threshold *after* the correct response). If the incorrect unit reached the predefined threshold of .10, the trial was classified as a partial error trial. In order to have enough trials available for analysis in both the simulation and the experiment, we focused on incompatible trials containing partial errors, since the number of partial errors is higher in incompatible situations (Burle *et al.*, 2002b; Coles

---

4. Botvinick *et al.* (2001) implemented two versions of the model : one with a feedback on the attentional layer, one without. The version used here implements the feedback.

*et al.*, 1985; Gratton *et al.*, 1988; Smid *et al.*, 1990).

The individual partial error trials were thereafter imported into EEGLAB (Delorme et Makeig, 2004) for further analysis (see “Single trial analysis” section).

### 3.2.2 Experimental study

#### Participants

Ten right-handed participants (3 women, 7 men, aged from 20 to 31 years, mean age = 25 years) volunteered for this experiment. They all had normal or corrected-to-normal vision. Before the study, all the participants gave their informed written consent according to the Declaration of Helsinki. They were informed of the purpose and procedure of the experiment before participating.

#### Experimental task and procedure

The participants performed an Eriksen’s flanker task (Eriksen et Eriksen, 1974) in which they had to respond with a right or a left thumb keypress as fast as possible as a function of a central target letter (S or H) flanked by two distractors that could be compatible (*e.g.* SSS) or incompatible (*e.g.* HSH). The stimuli were presented by a seven-segments light-emitting-diodes display (Lextronic, model SGN-S5,  $33 \times 14$  mm), located 1.5 m in front of the participant. The stimuli were extinguished with the participants’ response. The participants held vertical hand-grips on top of which response buttons were fixed.

Participants performed 20 blocks of 128 trials each. After stimulus presentation, they had 1 s to respond. The next stimulus was delivered 1 s after the response. All types of trials (HHH, HSH, SHS and SSS) were equiprobable and presented in a pseudo-random order.

#### Data acquisition and preprocessing

Electroencephalographic (EEG) and electromyographic (EMG) activity was recorded with Ag/AgCl electrodes (BIOSEMI Active-Two electrodes, Amsterdam). The sampling rate was 1024 Hz (filters : DC to 268 Hz, 3 dB/octave). For EEG, we used 64 channels (10-20 system positions). The vertical EOG was recorded by means of two electrodes (same type as EEG) just above and below the left eye, respectively, and the horizontal EOG was recorded with two electrodes positioned over the two outer canthi. EMG was recorded by means of two pairs of electrodes glued to the skin of the thenar eminence above the *flexor pollicis brevis* of each hand. The distance between the two EMG electrodes was 2 cm.

After acquisition, the electrophysiological data were filtered (EEG : high-pass = 0.3 Hz, low-pass = 100 Hz and EMG : high-pass = 10 Hz). Eye movement artifacts were corrected by the statistical method of Gratton *et al.* (1983). All other artifacts were rejected after visual inspection of individual traces. The onset of the EMG activity was marked manually after visual inspection. Indeed, although automated algorithms can be useful, visual inspection remains the most accurate technique against which all algorithms are compared (Staude *et al.*, 2001; Van Boxtel *et al.*, 1993), especially for detecting small changes in EMG activity such as partial errors. Importantly, the experimenters were not aware of the nature of the trial (compatible *vs* incompatible) being processed. Furthermore, the EEG signals corresponding to the current EMG were not displayed when detecting EMG onset. Thus, the experimenter was completely blind regarding all the other relevant parameters, and hence could not, even unwittingly, bias the results.

The trials were classified as correct or erroneous, depending on whether the correct or the incorrect button was pressed first. Among the correct trials, we separated trials containing only one EMG activation on the correct side, from trials containing an EMG activation on the incorrect side preceding the correct response (partial error trials, see Figure 3.1A). Laplacian transformation, as implemented in BrainAnalyser<sup>©</sup> (Brain Products, Munich), was applied to each individual trial to increase the spatial resolution of the EEG (Babiloni *et al.*, 2001) : first the signal was interpolated with the spherical spline interpolation procedure (Perrin *et al.*, 1989), and hence the second derivatives in two dimensions of space were computed. We choose 3 for the degree of spline since this value minimizes errors (Perrin *et al.*, 1987), and the interpolation was computed with a maximum of 15 degrees for the Legendre polynomial. We assumed a radius of 10 cm for the sphere representing the head, rather than the unrealistic default radius of 1 m assumed by BrainAnalyser<sup>©</sup>. With such a realistic radius, the most suitable unit is  $\mu V/cm^2$ . The individual Laplacian transformed trials were imported in EEGLAB for further analysis.

### 3.2.3 Temporal overlap estimation

In order to compare temporal overlap in simulated and experimental data, we choose a functionally equivalent measure for both. This measure, the  $\Delta$  value depicted on Figure 3.1, corresponds to the time between the onset of the partial error and the moment of the correct response (Figure 3.1A) for the experimental data, and to the “time” (represented in “cycles”) between the partial error threshold (.10) and the correct response one (.18, Figure 3.1B) for the simulations. In both cases, a greater  $\Delta$  value indicates a longer time between the incorrect response activation and the

correct response, revealing a lower temporal overlap. In the following we will use the  $\Delta$  value as a measure of temporal overlap.

### 3.2.4 Single trial analysis

EEG signal analysis normally relies on averaging techniques. Averaging, however, induces a considerable loss in the dynamics of the process of interest (Jung *et al.*, 2001), as will also be exemplified below. We therefore resorted to the “ERP-image” technique, implemented in the EEGLAB software (Delorme et Makeig, 2004), allowing one to visualize brain activity without averaging. This technique has been detailed elsewhere (Jung *et al.*, 2001), and will only be briefly described here. Note that this technique was applied not only to the Laplacian-transformed EEG data, but also to the simulated data. To construct the individual ERP-images, the trials are first sorted based on a relevant measure. In our case, for both the simulation and the experimental data, the time-zero corresponds to the onset of the partial error, and the trials are arranged by increasing  $\Delta$  values. The trials are then plotted as parallel colored lines. The result is a “raster-like” plot, with the x-axis representing time, the y-axis representing the arranged trials, and a color code indicating the intensity of the signal for each trial and each time point. On all figures, the vertical black line indicates the onset of the partial error, and the S-shaped one indicates the correct response. Below each raster-like plot, the average of the traces are represented as a function of time, hence giving an estimate of the activity under analysis. For the simulation, the blue color indicates conflict (positive values), whereas for the experimental data, blue represents negative polarity. This was done to improve the comparability between the conflict and the  $N_E$ , both appearing in blue in the ERP-images.

ERP-images have been computed for each participants : for the experimental data, the ERP-images are based on the individual Laplacian-transformed trials. For the simulation, this was performed on the individual activation function and/or the computed conflict traces. Besides individual subject representations, we sought for a population-based representation of the data. One way of doing so, is to put all the trials of all the subjects in the same ERP-Image. This approach has some strengths and limitations. The main interest is that one can visualize all trials of all subjects. However, the within-subject variance, in amplitude and in  $\Delta$  values, is confounded with the between-subject variance : for example, subjects with short  $\Delta$  values will mainly contribute to the lower part of the ERP-Image, whereas subjects with long  $\Delta$  values will mainly be present on the upper part. A similar bias holds for amplitude. Finally, data represented in this way are

not comparable to more traditional grand-averaged representation. One alternative approach is to build a “grand averaged” ERP-image by averaging the individual ERP-Images (see Appendix A for further details). In this case, all subjects have the same weight at every points of the ERP-Image, allowing one to visualize the impact of  $\Delta$  values on the population without bias induced by between-subject variability. One disadvantage, however, is that we do not see in this case “real” individual trials, but averaged ones for a given normalized  $\Delta$  value. Since these two methods have complementary strengths and shortcomings, we present the two representations along with individual ERP-Images. Since these various approaches led to very consistent results, despite their different strengths and drawbacks, we considered that the features expressed on these analyses could not be artifacts induced by any of the methods.

As recommended by Jung *et al.* (2001) the ERP-images have been smoothed, with a smoothing width set at about 10% of the number of trials. One exception, however, is to be noted for the primary motor cortices activations (Figure 3.7C and D) : since the activities of interest were of small amplitude leading to a lower signal/noise ratio, the smoothing step was set at about 20%.

### Statistical analysis

ERP-images, although very informative, do not allow one to statistically validate the observed features. To do so, both the simulated and experimental trials were binned into different classes depending on the  $\Delta$  interval. The classes were of equal width.

For the simulation, four classes of 3 cycles were retained that contained the largest number of trials for a reliable estimate of the conflict amplitude : class 1 = from 2 to 4 cycles, class 2 = from 5 to 7 cycles, class 3 = from 8 to 10 cycles and class 4 = from 11 to 13 cycles. The conflict signals obtained on each trial were then averaged, time-locked to the partial error onset for each class, and each run separately. The peak of conflict was determined as the maximum value in a window starting from the partial error onset and lasting 20 simulation cycles. We measured the peak and latency of this peak. Besides these “static” parameters, we investigated the dynamic aspects of conflict by studying how conflict develops. To do so, we analyzed the rising slope of the conflict by fitting a linear regression to the conflict signal in a window from the partial error onset to the 4<sup>th</sup> cycle following the partial error onset.

For the EEG data, the same method was applied : individual trials were binned as a function of the time separating the incorrect EMG onset from the correct response. Four classes that contained the larger number of trials (class 1 : from 101 to 150 ms, class 2 : from 151 to 200 ms, class 3 : from 201 to 250 and class 4 : from 251 to 300 ms) were retained. The EEG activity was averaged

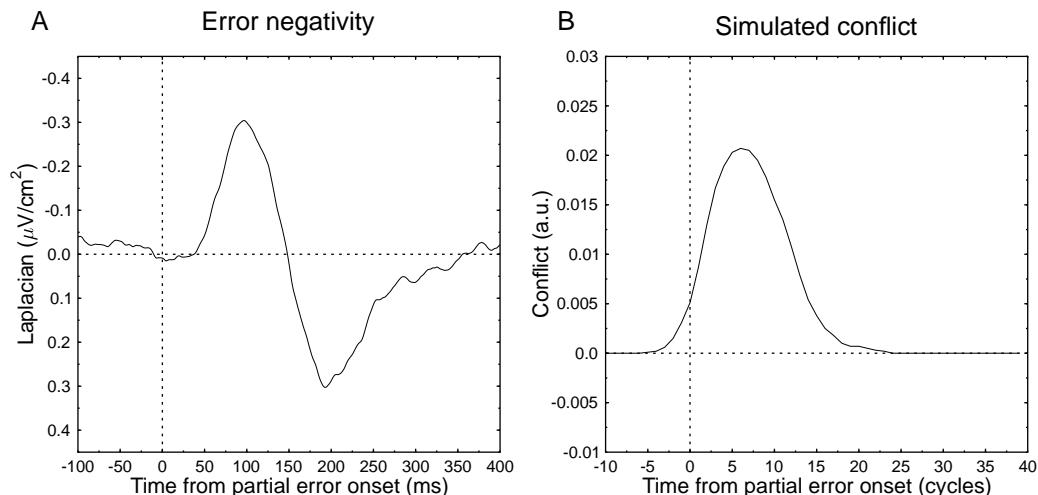


FIGURE 3.2 – **Comparison between grand-averaged  $N_E$  and conflict.** (A) Amplitude of the Laplacian transformed  $N_E$  induced by partial errors as a function of time. The time-zero corresponds to the partial error onset. A clear  $N_E$  can be observed starting about 30 ms after incorrect EMG onset and peaking about 100 ms after it. Furthermore, its topography is well localized fronto-centrally (see Figure 3.8 for the topography of this wave). (B) grand-averaged conflict as a function of time. Time-zero corresponds to crossing of the partial error threshold. The conflict is clearly maximal just after the partial error, and its timing nicely fits that of the  $N_E$ . This confirms that the  $N_E$  on partial errors is a valid measure for testing the conflict model predictions.

for each participant and for each class separately. As for the simulation, the parameter used to estimate the  $N_E$  was the amplitude of the peak of the  $N_E$  (defined as the difference between the positive peak occurring just after the partial error – between 10 and 50 ms – and the following negative peak – between 50 and 150 ms). We also analyzed how the  $N_E$  develops in time, by fitting a linear regression on the rising slope in a time window from 50 to 100 ms.

The statistical analysis involved either Student  $t$  tests, for comparisons between two means, or ANOVAs, for comparisons of more than two means. When ANOVAs were performed, the error term was always the interaction between the factor “participants” and the factor under analysis. Percentages and rates can not be tested directly with parametric methods, since their mean and variances are closely related. However, the arcsine transform ( $p' = \text{asin}\sqrt{p}$ , with  $p$  being the rate under analysis) has proved to be efficient in stabilizing the variances (Winer, 1971), and was therefore used consistently for each analysis involving rates.

### 3.3 Results

#### 3.3.1 Behavioral data

The overall percentage of errors was 5.2%. The number of overt errors was higher in the incompatible condition (7.5%) than in the compatible one (2.9%),  $t_9 = 6.55; p < .001$ . For correct trials, reaction time was longer when the flankers were incompatible (416 ms) than when they were compatible (386 ms),  $t_9 = 15.36; p < .001$ .

The number of partial error trials (see Figure 3.1) was higher in the incompatible situation (21.7%) than in the compatible one (14.3%),  $t_9 = 5.78; p < .001$ , in line with previously reported empirical data (Burle *et al.*, 2002b; Coles *et al.*, 1985; Smid *et al.*, 1990) (more behavioral data, and a deeper comparison between the experimental and the simulated results on these issues are presented in Appendix B).

#### 3.3.2 Comparison between conflict and $N_E$ timing as a function of $\Delta$

In previous simulation studies (Botvinick *et al.*, 2001; Yeung *et al.*, 2004), conflict was analyzed time-locked to the stimulus and/or to the response. Since here we introduce partial errors to the simulation, we first verified that a peak of conflict was indeed obtained just after the partial error. Figure 3.2A presents the grand-averaged  $N_E$  time-locked to partial error onset. As already reported (Allain *et al.*, 2004a; Burle *et al.*, 2005; Masaki et Segalowitz, 2004; Vidal *et al.*, 2000), a clear  $N_E$  was apparent just after partial errors. Furthermore, the  $N_E$  observed after partial errors is clearly localized fronto-centrally (see Figure 3.8). Figure 3.2B shows the grand-averaged conflict obtained in the simulations when time-locked to the partial error. As anticipated, the conflict is maximal just after the partial error, validating the comparison between  $N_E$  and conflict on partial error trials.

Figure 3.3 show the single trial dynamics of the  $N_E$  and of conflict. Panels A and C present the grand-averaged ERP-Images, whereas panels B and D present all trials of all subjects plotted together. Pannels E, F and G present single subject ERP-Images (see above for a presentation of these different approaches). Time-zero corresponds to the partial error onset in all cases, and the S-shaped black line indicates the occurrence of the correct response. One striking feature that appears in the comparison of the two graphs, is the difference in timing of the conflict and the  $N_E$  : although the simulated conflict shows an S-shape very similar to the correct response one, the  $N_E$  seems better time-locked to the partial error onset. It seems, however, that the width of

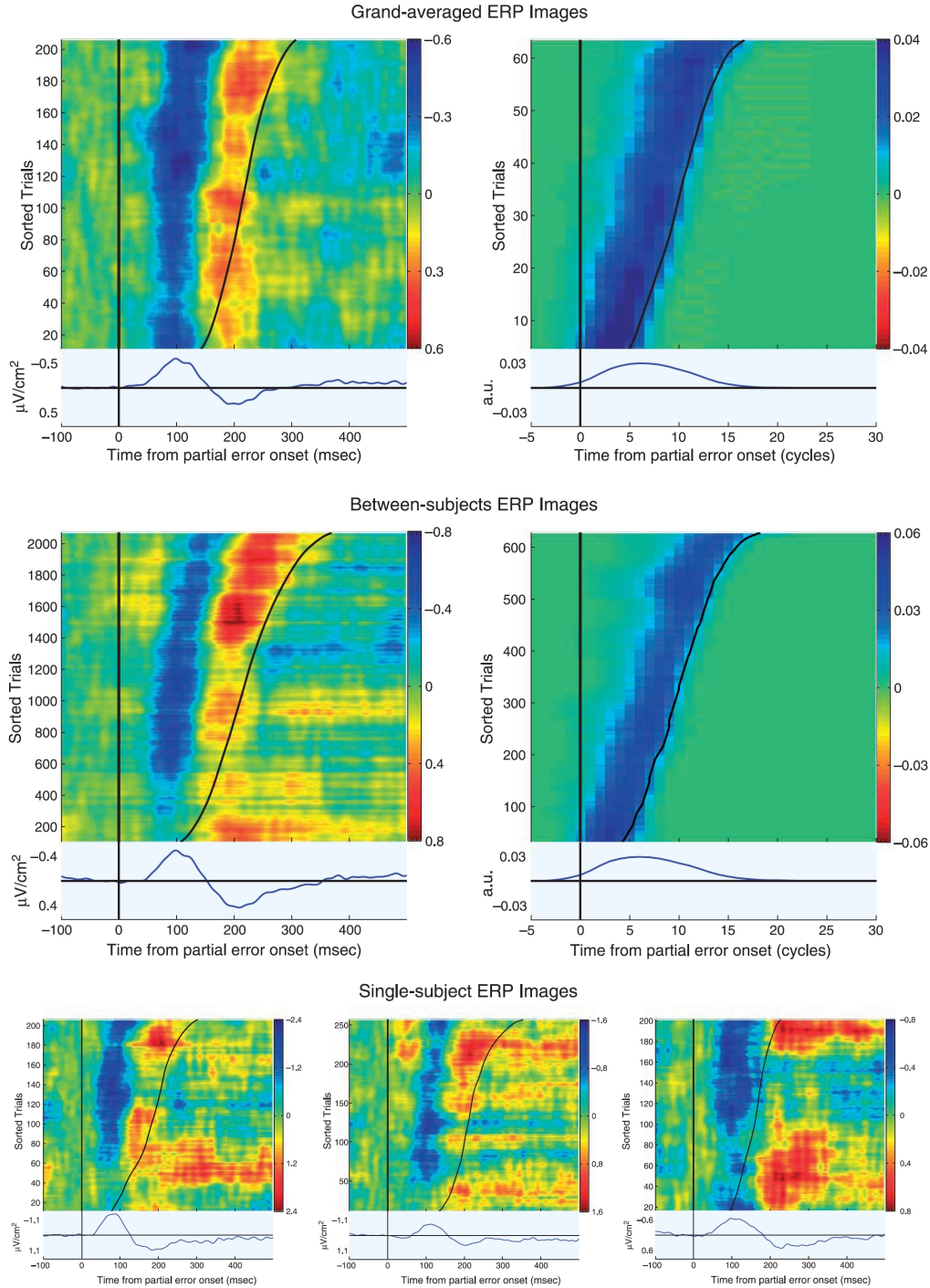


FIGURE 3.3 – **ERP-Images of the  $N_E$  and of the simulated conflict.** The trials containing a partial error were sorted as a function of increasing  $\Delta$  values. The vertical black line indicates the onset of the partial error and the S-shaped black line indicates the moment of the correct response. (A and B) Grand-Average ERP-Images of  $N_E$  and simulated conflict. (C and D) Grand ERP-Image, containing all trials of all “subjects” for the  $N_E$  and the simulated conflict, respectively. (E,F and G) examples of individual ERP-Images. Whatever the representation, the conflict appears more time-locked to the correct response, contrary to the  $N_E$ , which appears more time-locked to the incorrect response activation.



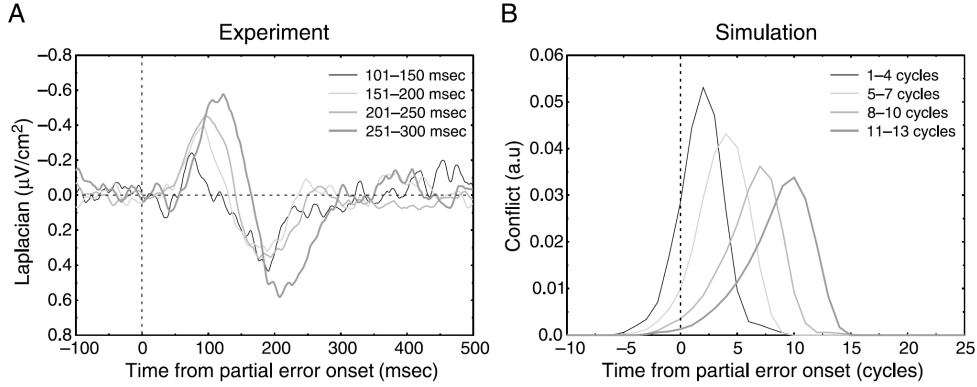


FIGURE 3.4 – (A) Classes analysis : Amplitude of the  $N_E$  (30 Hz low-pass filtered) as a function of time, for the four classes retained. Time zero corresponds to the onset of the partial error (see Figure 3.1). The amplitude of the  $N_E$  increases as  $\Delta$  increases. (B) Amount of conflict as a function of time, for the four classes retained. Time zero corresponds to the moment at which the partial error threshold (.10) was crossed. The amplitude of the conflict, measured at the peak, decreases as  $\Delta$  increases, therefore invalidating the proposition that the  $N_E$  reflects conflict.

the  $N_E$  increases as  $\Delta$  increases. To clarify these points, we binned the trials into different classes varying in term of  $\Delta$  (see the “Methods” section).

The grand average EEG data for each class is presented on Figure 3.4A. An ANOVA conducted on the peak amplitude revealed an effect of  $\Delta$ ,  $F(3, 27) = 3.91, p < .05$  : the smaller the  $\Delta$ , the smaller the  $N_E$ . The latency of the peak was also affected by  $\Delta$ ,  $F(3, 27) = 7.42, p < .001$ , with the peak occurring earlier for small  $\Delta$  than for large  $\Delta$ . In contrast, no effect of  $\Delta$  was observed on the rising slope of the  $N_E$ ,  $F(3, 27) = 1.34, p = .28$ . Therefore, the latency effect is simply a consequence of the fact that, with the rising slope being the same but the amplitude higher, the peak is reached later.

The same analysis was performed on conflict (Figure 3.4B). ANOVAs revealed a clear effect of  $\Delta$  on the amplitude of conflict,  $F(3, 27) = 400.24, p < .001$  : the smaller the  $\Delta$ , the higher the peak of conflict. The latency of the peak of conflict was also affected by  $\Delta$ ,  $F(3, 27) = 409.17, p < .001$ , with a peak of conflict occurring later as  $\Delta$  increased. Finally, the rising slope of the conflict was also steeper when  $\Delta$  was low than when it was high,  $F(3, 27) = 20.72, p < .001$ . It therefore appears that the amplitude of conflict decreases as  $\Delta$  increases. The dynamics of the conflict are also affected, since the rising slope is sensitive to  $\Delta$ .

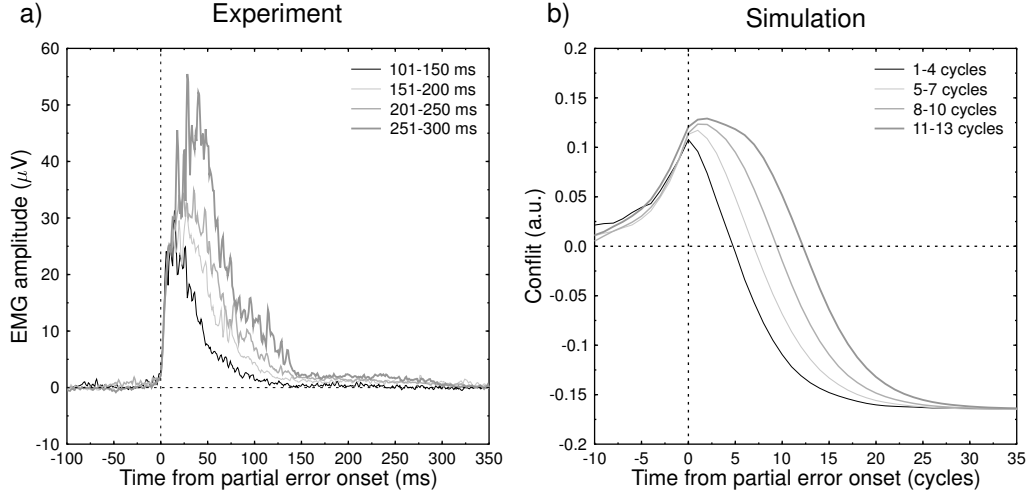


FIGURE 3.5 – **Amount of incorrect response activation as a function of  $\Delta$ .** (A) Grand averaged full-wave rectified EMG for the four classes retained. The amount of EMG increases as  $\Delta$  increases. (B) Grand average of incorrect response activation in the simulation, for the four classes. As for EMG, the amount of incorrect response activation increases as  $\Delta$  increases. The horizontal line at .10 indicates the threshold for a partial error.

### Alternative measure of conflict

Although the sensitivity of the  $N_E$  to  $\Delta$  is opposite to that of conflict as defined in the model, conflict seems to last longer when  $\Delta$  increases. We therefore evaluated whether a slight modification of the conflict computation could account for the data : instead of assuming that conflict at time  $t$  is the degree of co-activation of the two responses at time  $t$ , let's consider that it reflects the integration of such co-activation over time, that is conflict at time  $t$  could be :

$$Co(t) = \sum_0^t -2 \times (act_H(t) \times act_S(t) \times -3) \quad (3.3)$$

As a matter of fact, this quantity is assumed to be used for sequential adjustments (Botvinick *et al.*, 2001, see however Burle *et al.*, 2005). We therefore estimated the surface under the conflict curve for each simulated participant and for each class. This analysis revealed no change in conflict as a function of class ( $F < 1$ ). Therefore, even the integrated conflict does not increase as  $\Delta$  decreases, contrary to the  $N_E$ .

### 3.3.3 Conflict and the dynamics of incorrect and correct response activation

The above presented results are in clear disagreement with the interpretation of the  $N_E$  in terms of conflict. More specifically, the amplitude of the  $N_E$  does not parallel the degree of conflict. To better understand where the discrepancy comes from, we studied in more detail the activation dynamics of the incorrect and correct responses.

#### Conflict and the amount of incorrect response activation

When referring to eq. 3.2, it appears that conflict does not only depend on the temporal overlap between the two responses, but also on the amount of incorrect response activation (see simulation 5 of Yeung *et al.*, 2004). Differences in incorrect response activations might explain the observed discrepancy. To evaluate this, we averaged the simulated activation function of the incorrect responses, time-locked to the partial error onset for the four classes retained above. We processed the experimental data in a similar way : we took the size of the partial error EMG burst as an index of the incorrect response activation. To do so, the EMG bursts were rectified and then averaged separately for the same four classes used for the  $N_E$ . Figure 3.5 presents the results of this analysis. As one can see, for both the experimental (Figure 3.5A) and simulated (Figure 3.5B) data, the amount of incorrect response activation was greater when  $\Delta$  was high. Indeed, the size of the partial error (*i.e.* the amount of EMG) increased as the interval between the incorrect and correct response activation got larger,  $F(3, 27) = 24.43, p < .001$ . The linear component was also significant,  $F(1, 9) = 46.67, p < .001$ . The simulation results were similar : the amount of incorrect response activation increases as  $\Delta$  increases,  $F(3, 27) = 71.45, p < .001$  ; linear component :  $F(1, 9) = 160.1, p < .001$ . These results have two important consequences. The first one is that, as anticipated, conflict is more sensitive to temporal overlap than to the mere amount of incorrect response activation. Indeed, the maximum conflict was obtained in the situation where the activation of the incorrect response was the lowest. Second, since the same increase in incorrect response activation is observed in the experimental and simulated data, the difference between the  $N_E$  amplitude and conflict cannot be explained by a difference in response activation.

#### Dynamics of correct and incorrect responses activations

We further explored the discrepancy between conflict and the  $N_E$ , by investigating the “empirical conflict” present in the experimental data, that is the degree of co-activation of the alternative

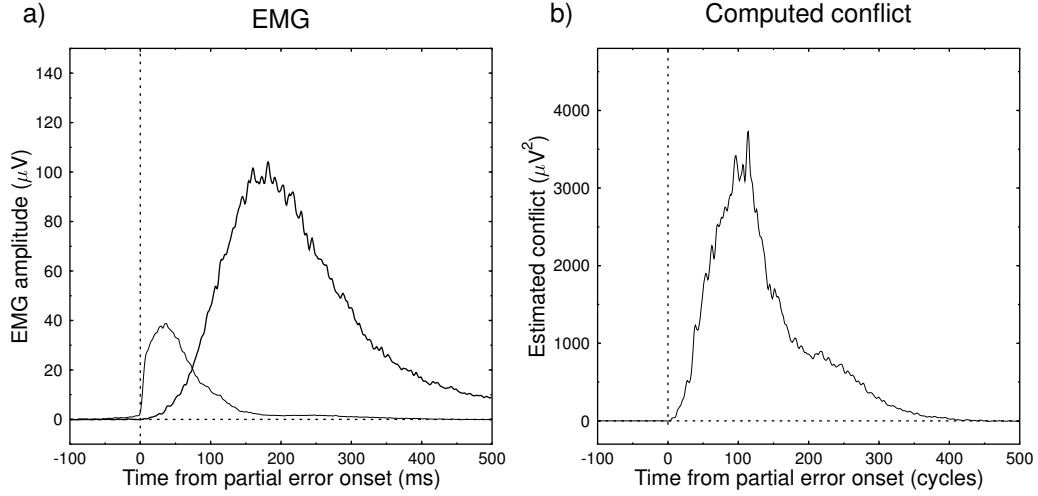


FIGURE 3.6 – **Empirical co-activation and conflict measured on the averaged traces.** (A) Temporal overlap of incorrect and correct response activation as measured with averaged EMG. The two responses appear to be co-activated around 100 ms after partial error onset. (B) empirical conflict obtained by applying eq. 3.2 to the averaged EMG activities. The empirical conflict nicely fits the prediction of the model, with conflict occurring shortly after the incorrect response activation, and peaking around 100 ms. This similarity, however, is an artifact due to averaging.

responses. We first analyzed the dynamics of response activation as estimated at the EMG level. Indeed, EMG seems a consensual measure of response activation, since, according to Yeung *et al.* (2004, p. 937) “Gehring et Fencsik (1999) have reported [...] that the ERN coincides with periods of co-activation of the correct and incorrect responses as measured through electromyography (EMG)”. A similar argument was also put forward by Botvinick *et al.* (2001, p. 635). To evaluate the “empirical conflict” we applied eq. 3.2 to the averaged EMG traces, time-locked to the incorrect EMG onset (Figure 3.6A).

The two *averaged* EMG activities overlap around 100 ms after partial error onset, which replicates previous results (Gehring et Fencsik, 1999; Masaki et Segalowitz, 2004). Figure 3.6B presents the estimated conflict, after applying eq. 3.2 to these *averaged* EMG data. The estimated “conflict” starts just after the partial error onset, and peaks about 100 ms after it. This timing almost perfectly fits the one of the  $N_E$  reported on Figure 3.2A. This replicates the data of Gehring et Fencsik (1999). This apparent similarity is, however, an averaging artifact (Callaway *et al.*, 1984), as demonstrated below.

The first row of Figure 3.7 presents the ERP-images for the incorrect (panel A) and correct (panel B) EMG. A striking effect emerges from these ERP-images : although the duration of the

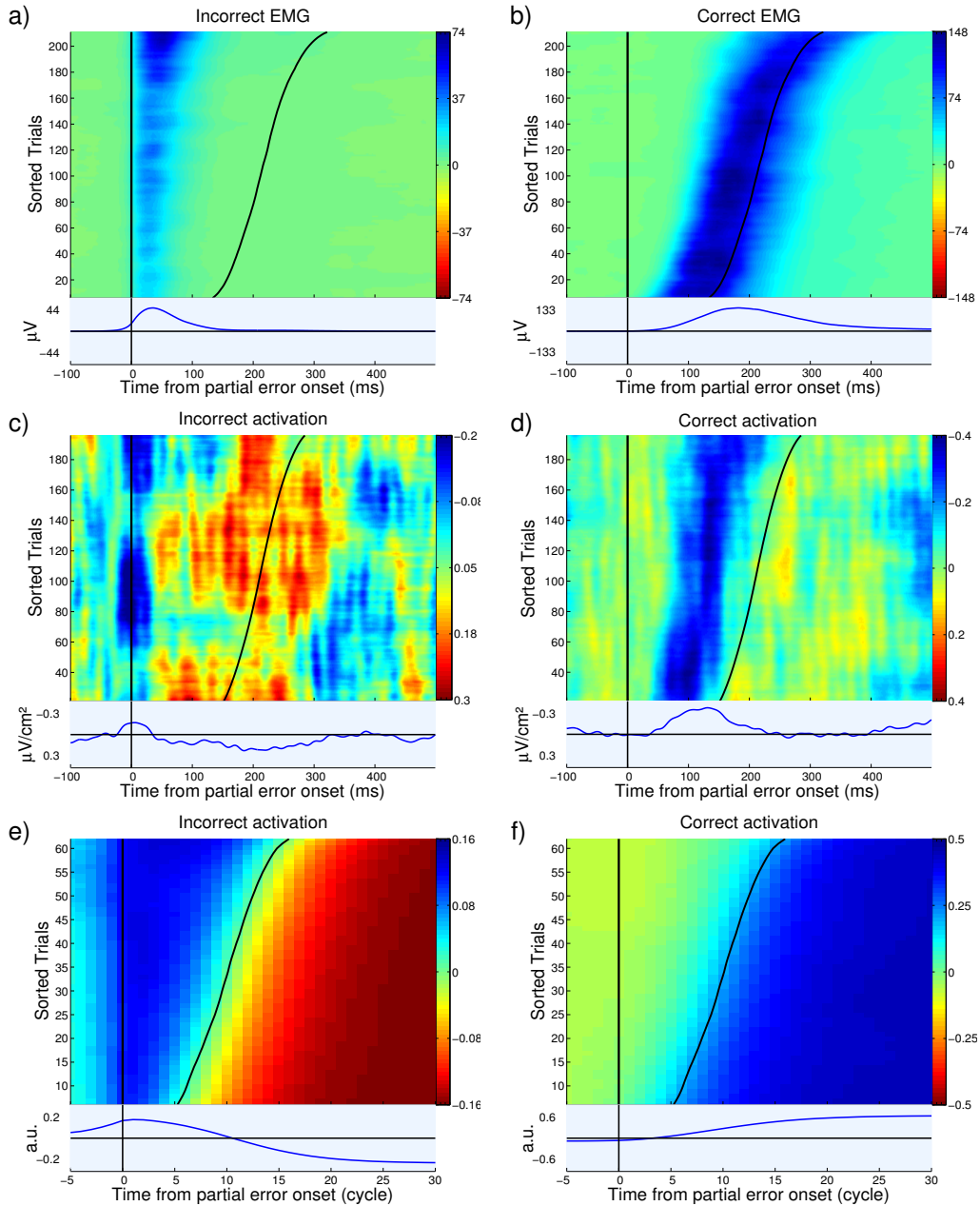


FIGURE 3.7 – **Single trial analysis of response activation.** This figure presents the incorrect (left column) and the correct (right column) response activations as a function of  $\Delta$ , for the experiment as estimated by EMG amplitude (first row), or at the cortical level through the Laplacian estimated over the primary motor cortices (second row), and for the simulation (third row)

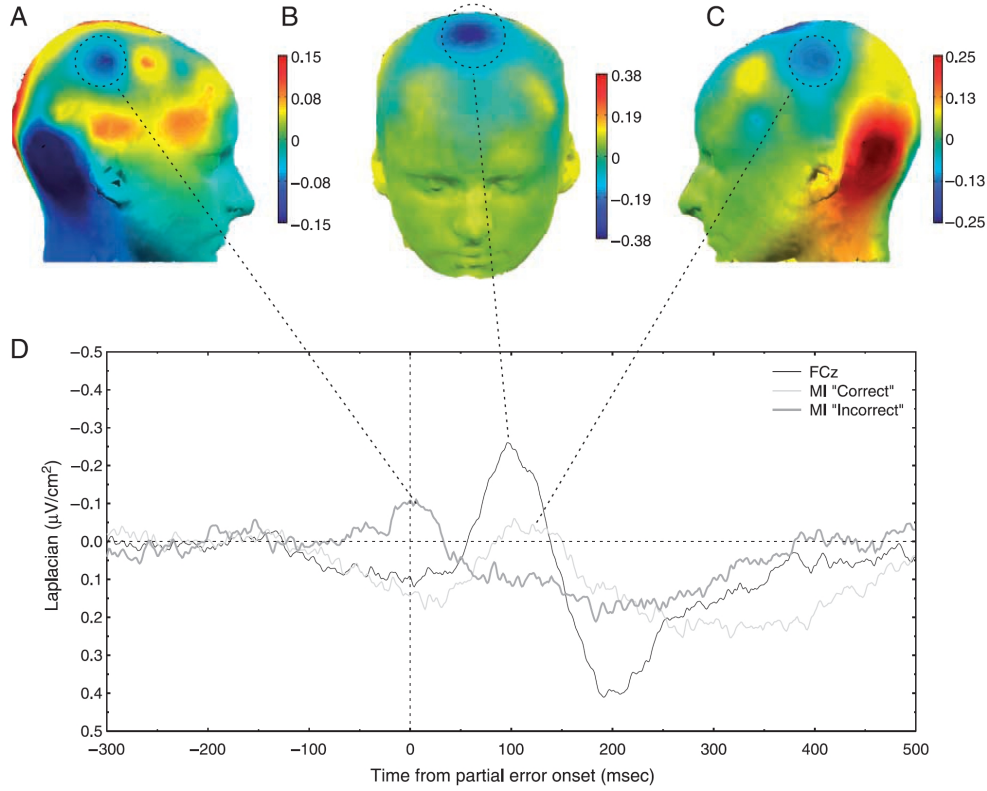


FIGURE 3.8 – **Topographies and time-courses of the activities under interest.** For sake of simplicity, the activities related to the two responses were averaged after mirroring the activities for the left response : activities recorded over the left electrodes for the were projected to their corresponding right ones, and vice-versa. Thus C3 corresponds to the primary motor cortices contralateral to the correct response, whereas C4 corresponds to the primary motor cortices contralateral to the incorrect response (*i.e.* the partial error. Except when noted, the baseline for the maps have been taken between -500 and -100 a) Laplacian (CSD) map obtained at incorrect activation time (9 ms after EMG onset). One observes a clear activity localized over C4, that is at the level of the primary motor cortex involved in the incorrect response. b) Laplacian map obtained at the peak of the  $N_E$ . The topography is clearly fronto-central, compatible with source(s) in the ACC and/or in the SMA. c) Laplacian map at the time of the correct response activation. One can observe a clear activity nicely localized over C3, that is over the primary motor cortices contralateral to the correct response. Note that, because of the positivity starting just before partial error onset (likely reflecting an inhibition, see Burle et al., 2004), the baseline was taken between -50 and 50 ms.

EMG burst is longer when the interval between the incorrect and the correct response increases (see Figure 3.5), the degree of overlap between the two EMG activities is virtually null, except maybe when the interval is very small. Thus, as long as response activation is evaluated at the EMG level (Yeung *et al.*, 2004), the incorrect and correct response activations did not overlap in any of the trials, and the conflict is null in this dataset. The apparent overlap observed on Figure 3.6 stems from the fact that the *slowest* incorrect activations occur later than the *fastest* correct ones, but not on the same trials. This exemplifies the hazards of averaging procedures, since, although we do observe “conflict” on averaged data that very nicely fit the predictions of the model, there is no “conflict” (as defined in the model ...) at all, on any of the trials, in this dataset.

One may argue, however, that even if co-activation does not occur at the EMG level, it might well be present at the central level. We thus analyzed the activity of the primary motor cortices contralateral to the incorrect and correct responses. Indeed, thanks to Laplacian estimation, it recently became possible to estimate the activation of the two primary motor cortices separately in choice RT tasks (Praamstra et Seiss, 2005; Tandonnet *et al.*, 2003; Taniguchi *et al.*, 2001; Vidal *et al.*, 2003b, see Burle *et al.*, 2004b for an overview).

The second row of Figure 3.7 presents the cortical activation of the incorrect response, *i.e.* the activity of the primary motor cortices contralateral to the partial error (left column) and the cortical activation of the correct one, *i.e.* the activity of the primary motor cortices contralateral to the correct response. The topographies of these activities, along with that of the  $N_E$ , are presented in Figure 3.8.

The signal/noise ratio of EEG being much lower than that of EMG, the results have to be interpreted with some caution. Some systematic patterns emerge nonetheless : the incorrect response activation, as measured at the primary motor cortices contralateral to the incorrect response, starts just before incorrect EMG onset (blue strip on Figure 3.7C), peaks shortly after EMG onset and then decreases back to baseline, or slightly below. It seems that the incorrect response activation is over about 30-40 ms after incorrect EMG onset. The correct response activation (Figure 3.7D) follows the  $\Delta$  value and starts later as this value increases. More importantly, the activation of the correct response clearly starts after the end of the incorrect response activation, and there seems to be virtually no overlap between the incorrect and correct response activation. Thus, here again, no conflict, as defined in the model, seems to emerge at the primary motor cortices level. We shall return to this absence of “conflict” in the Discussion.

Another interesting aspect is to be noted : when comparing the dynamics of the incorrect response activation and of the  $N_E$  (see Figure 3.3A), it appears that the  $N_E$  invariably starts

around the end of the incorrect response activation. Since it starts *after* the end of the incorrect activation, the  $N_E$  cannot be responsible for the interruption of the incorrect response activation as this incorrect response activation had necessarily been detected earlier in order to be stopped. Thus, the  $N_E$  cannot reflect the detection of this incorrect response activation, invalidating a simple error-detection mechanism (see Gehring et Fencsik (1999) for similar arguments on EMG).

For the sake of comparison, we also plotted the ERP-images of the incorrect and correct response simulated activations. The third row of Figure 3.7 presents the single-trial dynamics of the incorrect (panel E) and correct response (panel F) activations obtained in the simulation. The incorrect response starts being activated just before the partial error onset, and lasts almost until the correct response. The incorrect response starts being de-activated when the correct response starts being activated. The conflict, depicted on Figure 3.3B, thus peaks just before the correct response, when the two responses are still activated and are maximally competing.

## 3.4 Discussion

In the last few years, the conflict-loop theory (Botvinick *et al.*, 2001, 2004, 1999; Carter *et al.*, 1998, 1999; Cohen *et al.*, 2000; Yeung *et al.*, 2004) has played a key role in cognitive control research since it proposed a unified model that aimed to account for both behavioral and neurophysiological data. The model introduced a very simple, though elegant, concept : response-conflict is defined as the degree of co-activation of the possible responses. It has been argued that monitoring conflict allows, without any “clever” cognitive system, to detect failures in ongoing processing, and to adjust subsequent behavior. The model explicitly relates conflict monitoring with ACC activity as observed with metabolic measures, and through specific EEG components. The main EEG component related to conflict is the  $N_E$ . We will first discuss the relationship between the  $N_E$  and conflict, and then discuss the very concept of conflict as a relevant notion for cognitive control.

### 3.4.1 $N_E$ and conflict

The  $N_E$  has played an essential role in the development of the conflict model (Botvinick *et al.*, 2001, 2004; Yeung *et al.*, 2004). Reciprocally, the conflict model has been extremely influential in the  $N_E$  literature, since a lot of studies have used the conflict model as a general framework for interpreting their results. However, although several studies have tested the conflict interpretation of ACC activity observed in metabolic research (Brown et Braver, 2005; Egner et Hirsch, 2005a), few studies have explicitly addressed the conflict model hypothesis on the  $N_E$ , apart from the conflict



modelers themselves (Botvinick *et al.*, 2001; Yeung *et al.*, 2004, see Carbonnell et Falkenstein, 2006 for an exception). Here, we directly tested the conflict interpretation of the  $N_E$  by estimating the amount of conflict on a trial-by-trial basis by analysing partial error trials (Figure 3.1) that allow, on a single trial basis, to track the activation of incorrect and correct responses. According to the definition of conflict, we reasoned that conflict should depend on the temporal overlap between the two response activations, that we measured through the  $\Delta$  value. A simulation confirmed that the amount of conflict depends on the  $\Delta$  between the two response activations : the smaller the  $\Delta$  between the incorrect and correct response activations, the larger the conflict. However, the present experiment shows that the  $N_E$  decreases as the  $\Delta$  decreases, therefore demonstrating that the  $N_E$  evolves in a way opposite to the conflict. The mandatory conclusion that derives from the present study is that the  $N_E$  does not reflect conflict. A note concerning the model's prediction regarding metabolic signal is in order. Indeed, it is important to note that fMRI, because of its low temporal resolution, provides information of a different nature compared to EEG : the measured signal represents a *temporal integration* of brain activity. In this respect, the BOLD signal should not be compared to the peak amplitude of the conflict, but more likely to the overall conflict, that is the surface under the conflict curve (*i.e.* the integral of the conflict over time). That is actually what Botvinick *et al.* (2001) did in their first simulations that sought to compare the conflict predictions with ACC activations obtained with fMRI. Integrated conflict was measured and analyzed in the section "Alternative measure of conflict", and the results show that the integrated conflict is the same whatever  $\Delta$ . Thus, interestingly, the model would predict no change in ACC BOLD response, as a function of  $\Delta$ , despite the fact that conflict is clearly affected by  $\Delta$ . This confirms that timing is essential for an appropriate test of the model.

The main hypotheses of the conflict model were : 1) the amount of conflict is monitored by the ACC, 2) the electrophysiological correlate of conflict monitoring is the  $N_E$ , 3) conflict triggers cognitive control which results in subsequent adjustments in behavior. The present data clearly invalidate prediction 2. Recently, prediction 3, namely that sequential effects depend on conflict, has been put into test by Burle *et al.* (2005). The results demonstrated that conflict, as assessed by incorrect response activation and reflected by the  $N_E$ , is *not sufficient* to account for sequential adjustments. Concerning prediction 1, one may argue that the ACC indeed measures conflict, but that the  $N_E$  does not reflect such a conflict monitoring. Although this considerably weakens the model, this view could save (a part of) it. This is, however, unlikely. Indeed, several source localization studies have pointed out that the  $N_E$  has a source in the ACC (Dehaene *et al.*, 1994; Van Veen et Carter, 2002), although other areas, including the Supplementary Motor Area, might

also contribute to the  $N_E$  (Dehaene *et al.*, 1994; Herrmann *et al.*, 2004; Stemmer *et al.*, 2004). Furthermore, recent co-registration of EEG and fMRI provided a strong argument in favor of the idea that the ACC is at least involved in the genesis of the  $N_E$  (Debener *et al.*, 2005). Hence, ACC activity, as assessed by fMRI, and  $N_E$  are likely to be strongly related.

A recent report, in light of other studies, further suggests that ACC does not monitor response conflict : di Pellegrino *et al.* (2007) showed that ACC is *necessary* for sequential adjustments to occur (see also Kerns *et al.* (2004)), while Burle *et al.* (2005) showed that response conflict is *not sufficient* to trigger those adjustments. Taken together, those two studies clearly dissociate ACC and response conflict monitoring.

Considering the above arguments, the most parsimonious position is to assume that, although ACC seems clearly involved in detecting the need for more cognitive control, it does not do so through conflict monitoring (Egner et Hirsch, 2005a; Nakamura *et al.*, 2005).

### 3.4.2 Response co-activation, conflict and behavioral interference

Besides the interpretation of the  $N_E$  and ACC activity in terms of conflict, the present data also question the very notion of response conflict – or response competition – as a universal explanation for both behavioral interference and ACC activity. Indeed, the EMG data clearly indicate that there was no co-activation of the responses at such a peripheral level, and the same conclusion seems to hold also at the level of primary motor cortex. Even for trials in which the incorrect response was undoubtedly activated, the correct and incorrect responses are never activated at the same time, casting some doubt on the general idea that the responses are competing. Before we go further in this direction, a comment is in order : the fact that there was no co-activation in the present dataset does not imply that co-activation is never obtained. As a matter of fact, Carbonnell et Falkenstein (2006) did observe overlap between response force traces on a trial-by-trial basis. Thus the presence of overlap between response activations might well depend on the specific parameters of the task. In any cases, the important point is that, even without co-activation, the present dataset clearly show an interference effect on RT. Thus, if co-activated responses compete, inducing an interference, the present data show that *response-competition* is not necessary for interference to occur. Co-activation seem also lacking at the primary motor cortices level. Note that, although they are not the only areas involved in actual response activation, there is general agreement that MI plays an essential role in the implementation of the motor command. Thus, if behavioral interference is due to a competition between mutually exclusive representations, such a competition does not

occur at the motor execution level (Burle *et al.*, 2002b; Rösler et Finger, 1993) but more likely upstream in the information processing chain, contrary to what is often assumed (Coles *et al.*, 1985; Gratton *et al.*, 1988, see however Valle-Inclán et Redondo, 1998). This may sound at odds with a rather large amount of data, empirical and theoretical, suggesting that response-competition is at the core of interference effects. However, it is to be noted that (almost) all arguments for motor co-activation come from EEG data, employing averaging procedures. Importantly, our averaged data also seem, at first sight, in agreement with motor co-activation (see Figure 3.6). However, as described above, the apparent co-activation stems from an artifact introduced by averaging and when one looks at the same data with methods that avoid the production of such an artifact, the observed pattern leads to the opposite conclusion. One may thus wonder whether similar artifacts are present in the literature (see Meyer *et al.* (1988) for similar concerns) and the arguments for co-activation of motor components may need to be re-evaluated.

Note, however, that the absence of co-activation of responses does not preclude the possibility of co-activation of mutually exclusive representations upstream from the primary motor cortices. Indeed, competition might well occur at more abstract (or central) levels. In this case, this would mean that interference does not occur at the motor level (Burle *et al.*, 2002b; Rösler et Finger, 1993; Valle-Inclán et Redondo, 1998), but at other stages of information processing. Thus, the present results might remain compatible with “competition” views, but push the location of such a competition upstream in the information processing chain.

### 3.4.3 Relations to other models of ACC/ $N_E$ function

The present data clearly invalidate the interpretation of the  $N_E$  in terms of conflict, defined as the co-activation of two responses. Some alternative models of ACC function and of the  $N_E$  have been proposed recently. We shall now discuss the implication of the present data for these models.

#### The Reinforcement–Learning Theory of the $N_E$

Holroyd and colleagues Holroyd et Coles (2002); Holroyd *et al.* (2005) proposed an alternative formal model of the  $N_E$  and ACC function<sup>5</sup>. This model is based on reinforcement–learning theory Sutton et Barto (1998) that implements the so-called *temporal difference error*. It is beyond our

---

5. Brown et Braver (2005) proposed a modified version of Holroyd et Coles (2002)’s model : the so-called “Error–likelihood” model. It aims mainly at accounting for metabolic data, and a recent test of the model for electrophysiological data invalidate some of its main predictions Nieuwenhuis *et al.* (2007). We will thus not further discuss this model here.

scope to describe the model in detail. We will briefly present the features that are relevant for evaluating the impact of the present results for this model. The most recent version of the model Holroyd *et al.* (2005) for (a modified version of) the Eriksen task is made of two components : a task module and a monitor module. The task module implements the operations necessary to solve the task. It is composed of three layers : an input layer (coding for the letters presented at each position – *i.e.* “H” on left position, or “S” on the center position etc...), a category layer (representing a decision concerning the nature of the central – *i.e.* the target – letter) and a response layer (representing the two possible responses)<sup>6</sup>. The category and the responses layers continuously send their level of activity to the monitor module. The monitor module is also composed of several layers, but we will present only the relevant ones. One set of units, so called *conjunction units* receive activation from the category and response units and code for the conjunction of the two (*i.e.* the *HL* unit receives activation from the stimulus unit *H* and the response unit *L*. Its activation thus indicates that the target *H* and the response *L* have been activated). There are thus four conjunction units (2 targets  $\times$  2 responses), and each of these units is associated with a *value*, that can be positive if the conjunction corresponds to a correct response, or negative in the opposite case. For example, if the instruction is to give a left response when the target is “H”, the value of *HL* will be positive, whereas the value of *HR* (“H” target and right response) will be negative. A *Temporal difference* unit (*TD*) receives inputs from the conjunction units, integrates them and issues a temporal difference signal, that will be an error signal in the case of incorrect response. More formally, since no more than one unit can be active at any time (Holroyd *et al.*, 2005, p. 179), the activation  $a_{TD}^t$  of the *TD* at time  $t$  amounts to :

$$a_{TD}^t = S_t \times V_s \quad (3.4)$$

where  $S_t$  is the present state of the system (the level of activity of the conjunction unit, if any) and  $V_s$  is the value of this state (Note that this corresponds to basic definition of conditioning learning processes, see Sutton et Barto (1990)). The temporal difference signal at time  $t$ , whose amplitude corresponds to the  $N_E$ , is defined as :

$$\delta^t = a_{TD}^t - a_{TD}^{t-1} \quad (3.5)$$

that is the difference between two successive time steps (represented as cycles in the model).

Can this model account for the present data? Although the simulations reported by Holroyd *et al.* (2005) were only concerned with overt errors, one can easily assume that an error signal could

---

6. An attentional layer is also present in the task module, biasing stimulus and response processing. However, since the impact of the attentional layer is not relevant here, we will not discuss it further

be generated by partial errors<sup>7</sup>. Thus in principle, partial error could certainly induce an  $N_E$  in this model. Furthermore, according to the way the error signal is generated, it is likely that the predicted  $N_E$  would be much more time-locked to the partial error onset than would the conflict (cf. Figure 3.3B) and, hence, more similar to the experimental  $N_E$  (cf. Figure 3.3A), although some simulations are certainly needed to better confirm this hypothesis. However, without significant modification, the model does not seem able to account for the relationship between  $N_E$  amplitude and  $\Delta$  values. Indeed, as clearly stated by (Holroyd *et al.*, 2005, p. 178) “[...] the first unit of each pair [of task state units] to be activated remains active until the end of the trial. Critically, if the task module generates a second response following an error (an *error correction*), the response detection unit activated by the initial response remains active, and the response detection unit associated with the second response remains inactive. [...]”. Thus, even if the task module could correct an error, the monitor module would be blind to such a correction, and hence the error signal could not be sensitive to the timing of such a correction.

If we relax this constraint, one may wonder whether a modified version of the reinforcement learning hypothesis could account for the data. Given that the value of a given state is constant during a trial (even if it can vary during learning), it appears from eq. 3.4 that the evolution of the  $N_E$  will only depend on the time course of  $S_t$ . If one accepts that, after initial activation of an incorrect conjunction unit (triggering an error signal, represented as a negative TD signal, see Holroyd *et al.* (2005)), a correct conjunction unit can later be activated, the activation of this correct unit will trigger a positive TD signal that will counteract the negative one (*i.e.* interrupt the error signal). Interestingly, in this case, there will be a monotonic relationship between the end of the error signal and the correction : the later the correction, the longer the duration, and probably the greater the activation, of the error signal. This would correspond to the  $N_E$  results (Figure 3.3C). Obviously, such a possibility needs to be implemented and tested in future simulations.

#### 3.4.4 Online control of control ?

Besides the invalidation of the conflict account, the present results also provide insights that may help in deciphering the functional significance of the  $N_E$ .

In the present data set, the latency and dynamics of the  $N_E$  onset are independent of the interval between incorrect and correct response activations, suggesting that the  $N_E$  induced by the partial error initially develops in the same way whatever the timing of the correction. The  $N_E$ ,

---

7. In a way similar to that conducted for the conflict model.

however, lasts longer and reaches a higher amplitude when this interval increases (Figure 3.3 A and C). After this initial development, the  $N_E$  seems interrupted. This interruption correlates with the timing of the correct response activation : the later the correction, the later the interruption, in agreement with data obtained by Fiehler *et al.* (2005) who observed that the  $N_E$  on overt errors peaks later when the correction of the error is slow.

The  $N_E$  initial development is similar whatever the timing of the correction, suggesting that the  $N_E$  is *a priori* the same, and hence that the observed differences occur later in time. This indicates that the  $N_E$  can be modulated “online”, that is during the course of a trial. Furthermore, the  $N_E$  interruption being directly linked to the correction, it seems that the  $N_E$  is suppressed once the remediation process has started. This suggests that the ACC activity is used as an “alarm signal”, which lasts until remediation processes take place, making such an alarm signal useless. In this case, this would indicate that the need for control is also monitored and adjusted online, during the course of a trial. In the above discussion of the reinforcement-learning model of the  $N_E$ , we have seen some possible directions on how this model could implement such an idea. Whatever the exact nature of this signal, since it is highly dynamic and flexible, this opens new perspectives, and adds constraints on possible modeling of the evaluation processes.

## 3.5 APPENDIX A

### Single trial dynamics

In order for all the participants to have the same weight on all the “trials” of the ERP-Images, we have computed the ERP-images for each participant separately, and then averaged the ERP-Images. ERP-Images are in fact matrices of size `number_of_points`  $\times$  `number_of_trials`. The value of `number_of_trials`, however, is not constant across subjects, since it depends on the number of partial errors. Thus the matrices for the various subjects had not exactly the same size, precluding direct averaging of the matrices and hence of the ERP-Images. In EEGLAB, the `number_of_trials` value in the matrix can be reduced by applying an inter-trial smooth Jung *et al.* (2001). One can thus theoretically reduced the dimensions of all the matrices of all the subjects to the same value. This, however, necessitates large smoothing values for the subjects presenting a lot of trials, given the large differences in number of trials (see table 3.1). We therefore choose a balanced option, by both decreasing the dimension of the matrices containing the largest number of trials by applying an appropriate smooth implemented in EEGLAB, and by increasing the size of

TABLE 3.1 – Summary of the number of partial errors and of the smoothing/interpolation procedure applied to the individua data. The final number of “trials” (*i.e.* rows in the matrix) was set to 216, since i) two participants had this number of trials, and ii) it seemed a good compromise between smoothing (reducing dimension) and interpolation (increasing dimension).

Subject N°	Partial errors	Smooth	Interpolation	Final
1	216	-	-	216
2	92	-	124	216
3	236	20	-	216
4	146	-	70	216
5	279	63	-	216
6	268	52	-	216
7	173	-	43	216
8	277	61	-	216
9	267	51	-	216
10	216	-	-	216

the matrices presenting a low number of trials, by applying an interpolation of the image accross the trials. Note that the interpolation does not change anything to the image but its size. The information conveyed by the ERP-Image is thus kept constant, as exemplified in Fig. 3.9. The smoothing/interpolation procedure applied to each individual participant is presented on table 3.1.

## 3.6 APPENDIX B

### Chronometry of partial errors

In the experimental data, the latency of the partial error onset decreased from class 1 ( $\Delta$  values from 101 ms to 150 ms) to class 4 (from 251 to 300 ms),  $F(3, 27) = 21.2; p < .001$ , linear component,  $F(1, 9) = 34.05; p < .001$ . Symetrically, the latency of the correct response increased,  $F(3, 27) = 59.81; p < .001$ , linear component :  $F(1, 9) = 86.13; p < .001$ . Thus the increase in the  $\Delta$  values is due to both earlier partial errors and later correct responses (see Table 3.3). A similar pattern was obtained in the simulation : The latency of the partial error onset decreased from class 1 ( from 2 to 4 cycles) to class 4 (from 11 to 13 cycles),  $F(3, 27) = 77.97; p < .001$ , linear

TABLE 3.2 – Summary of the number of partial errors and of the smoothing/interpolation procedure applied to the simulation. The final number of “trials” (*i.e.* rows in the matrix) was set to 67.

Simulation N°	Partial errors	Smooth	Interpolation	Final
1	68	1	-	67
2	67	-	-	67
3	63	-	4	67
4	57	-	10	67
5	71	4	-	67
6	74	7	-	67
7	57	-	10	67
8	49	-	18	67
9	85	18	-	67
10	67	-	-	67

component,  $F(1, 9) = 188.12; p < .001$ . Symetrically, the latency of the correct response increased,  $F(3, 27) = 70.87; p < .001$ , linear component,  $F(1, 9) = 169.27; p < .001$  (see Table 3.4). Thus, as for the experimental data, the increase in the  $\Delta$  values is due to both earlier partial errors and later correct responses. More importantly, the partial error recorded experimentally and simulated behave in the same way, making the two highly comparable.

TABLE 3.3 – Mean chronometric indices (in ms) obtained in the experiment for the partial error trials.

	Partial error onset	Response Time	$\Delta$ value
class 1	277	410	133
class 2	260	440	180
class 3	246	481	235
class 4	237	505	267



TABLE 3.4 – Mean chronometric indices obtained in the simulation for the partial error trials. All the values are given in number of cycles.

	Partial error onset	Response Time	$\Delta$ value
class 1	6.21	9.79	3.58
class 2	4.42	10.61	6.18
class 3	2.94	12.12	9.18
class 4	2.14	14.13	11.99

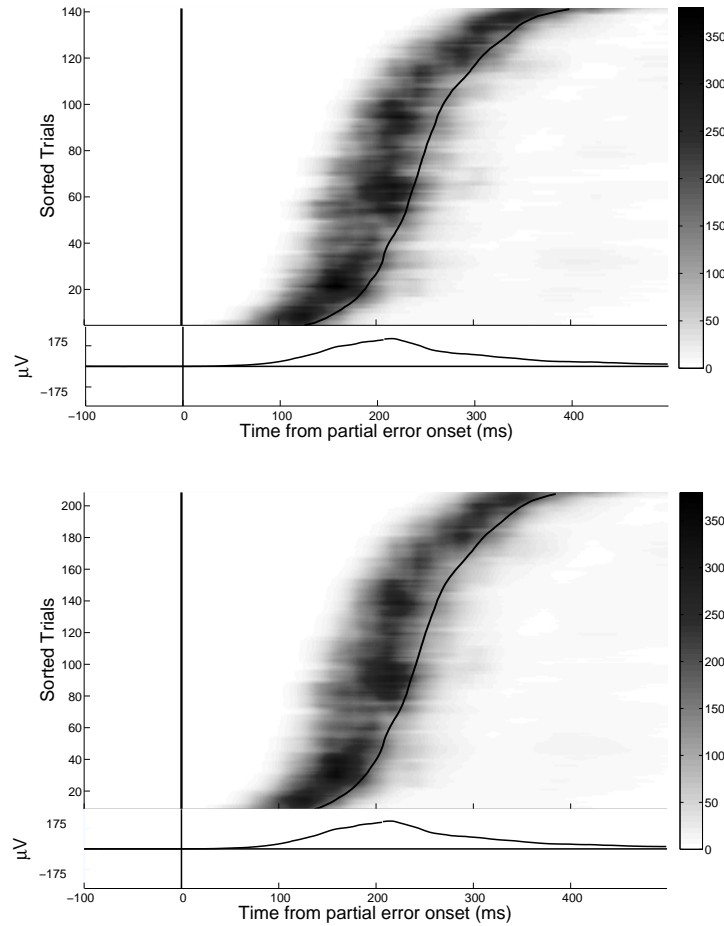


FIGURE 3.9 – **Impact of interpolation on the time-course of activity of interest** : The top figure represents the ERP image of EMG activity for subject N° 4, before interpolation (number of trials = 146), and the bottom figure represents the same data after interpolation (number of “trials” = 216). As one can see, the interpolation did not introduce any distortions nor remove any obvious properties of the signal. Colorbar are in  $\mu V$ .



## Chapitre 4

# Rostral Cingulate Zone and correct response monitoring : ICA and source localization evidences for the unicity of correct– and error–negativities<sup>1</sup>

### 4.1 Introduction

Keeping our behavior adapted to an ever changing environment requires a constant evaluation of one's own performance. In this respect, errors play an essential role in this evaluation, since they strongly signal the need for adaptation. In the early 90's, Falkenstein *et al.* (1991) reported the existence of an EEG component peaking just after error commission in reaction time (RT) tasks (see also Gehring *et al.*, 1993) : This fronto-central negative wave starts just before the response, and peaks between 50 and 100 ms later. With conventional monopolar recordings, this activity has originally been observed only on errors and was hence interpreted as reflecting an “Error Detection” mechanism. Accordingly, it was named “Error Negativity” (Ne, Falkenstein *et al.*, 1991) or “Error-

---

1. article en cours de révision dans le journal *NeuroImage*.

Related Negativity” (ERN, Gehring *et al.*, 1993). Source localization approaches (Dehaene *et al.*, 1994; Herrmann *et al.*, 2004; Van Veen et Carter, 2002) and fMRI data (Debener *et al.*, 2005; Ullsperger et von Cramon, 2001) have pointed to a Rostral Cingulate Zone (RCZ, Ridderinkhof *et al.*, 2004a) generator of this activity. This generator would be more likely located within the anterior cingulate cortex (ACC) and/or the supplementary motor area (SMA) (Dehaene *et al.*, 1994; Ullsperger et von Cramon, 2001). This Ne was later included in more general models of response monitoring, and was re-interpreted as reflecting conflict monitoring (Yeung *et al.*, 2004, see however Burle *et al.*, 2008a).

The specificity of the Ne to errors was disputed by Vidal *et al.* (2000). These authors computed the Current Source Density (by applying the Laplacian operator), which has been shown to dramatically improve the spatial resolution of monopolar recordings (Babiloni *et al.*, 2001). Thanks to this methodological improvement, Vidal *et al.* (2000) evidenced that a similar activity was also observed on correct trials, albeit with smaller amplitude. They first analyzed some particular correct trials in which *partial errors* occurred : On such trials, although the correct response was given, electromyographic (EMG) recordings allow to reveal a small EMG burst on the muscles involved in the incorrect response (Burle *et al.*, 2002b; Coles *et al.*, 1985; Eriksen *et al.*, 1985). Vidal and colleagues observed a negative wave just after the onset of such partial errors with comparable latency and topography as the wave reported on errors (see Scheffers *et al.*, 1996 for similar data in a go/nogo task). More importantly, they also reported a similar negativity, of smaller amplitude though, just after the EMG leading to the correct response on pure correct trials (*i.e.* trials without any sign of incorrect EMG activation).

The negative activities obtained on errors, partial errors and pure correct trials had similar topographies, similar time-courses (after Laplacian transform), and their amplitude was shown to monotonically decrease from errors to pure correct, with partial errors in between. Based on these similarities, Vidal *et al.* (2000) argued that the  $N_C$  was of same nature as the Ne on errors.

Coles *et al.* (2001) disputed this view and argued that the negativity reported by Vidal *et al.* (2000) on correct trials was due to an artifact caused by the temporal overlap between stimulus-locked and response-locked activities. To address this point, Vidal *et al.* (2003a) visualized the single-trial dynamics of the stimulus and response evoked potentials as a function of the reaction time, and showed that the  $N_C$  was clearly response-locked and independent from stimulus-locked activities. Several other studies have reported a negative wave in correct trials (Falkenstein *et al.*, 2000; Luu *et al.*, 2000b; Mathalon *et al.*, 2002), and there is now a consensus on its existence.

The question remains, however, to determine whether the negativities recorded on correct and

error trials reflect the same functional and physiological mechanisms modulated in amplitude or whether they are completely different processes.

While the proposition that the negativities observed on pure correct, partial errors and errors reflect the same, modulated, mechanism (Vidal *et al.*, 2000, 2003a), is supported by the fact that the negativity on correct trials is also sensitive to the subject's performance (Luu *et al.*, 2000b; Ridderinkhof *et al.*, 2003; Allain *et al.*, 2004c; Hajcak *et al.*, 2005b), this view was disputed by Yordanova *et al.* (2004). These authors reported that on correct trials the negativity tended to be lateralized toward the hemisphere contralateral to the responding hand whereas the topography was more central for errors. Based on the difference in topography and in the time-frequency pattern of negativities on correct and erroneous trials, they concluded that the two negativities reflect different processes. The lateralization reported by Yordanova *et al.* (2004) might well be due, however, to an independent source. Indeed, following the motor lateralization induced by response execution processes (Vidal *et al.*, 2003b, see Burle *et al.*, 2004b for a review), the lateralization of the  $N_C$  observed by Yordanova *et al.* (2004) could be due to the propagation of the primary motor activity towards premotor areas (see Tandonnet *et al.*, 2005, Fig. 1) : If this pre-motor activity is of same amplitude for correct and errors trials, it may contribute more to the topography when the amplitude of the medial activity is lower, that is for correct trials. This may give the false impression of a lateralization limited to correct trials, although the same lateralized activity could also be present on errors, but less visible. In agreement with this view, a critical look at the figures 1 and 6 of Yordanova *et al.* (2004) shows that, even on errors, the iso-contour lines present a lateralization.

Clarifying this debate is theoretically important since none of the current models of cognitive control can easily account for the presence of a " $N_C$ " on correct trials. Thus, if the negativity on correct trials were of the same nature as the negativity on errors, this would indicate that control processes operate gradually from correctness to errors. This would also open new perspectives and add new constraints on cognitive control modeling.

We therefore assessed the unicity of those negativities with Independent Component Analysis (ICA, Onton *et al.*, 2006) and source localization techniques (sLORETA, Pascual-Marqui, 2002).

Applied to EEG, ICA posits that the scalp activity is a linear combination of a limited set of elementary brain signals (the *independent components*). Based on the assumptions of temporal independence, ICA allows one to recover the mixture of components and hence to estimate the time course and the topography of each component. Of special interest in the present context, ICA *blindly* recovers the components, that is, it does so without any a priori assumptions about

the components (except that they must be maximally independent from each other). We reasoned that if the three negativities recorded on errors, partial errors and correct trials reflect the same modulated elementary brain activity, they should be captured by ICA in the same component. On the contrary, if they reflect different mechanisms, it would not be possible to find a single component accounting for these three waves.

In addition to the ICA argument, we also applied Source localization techniques to recover the generator(s) of the three activities. Indeed, although there is now strong arguments for an RCZ origin of the Ne on errors (Dehaene *et al.*, 1994; Debener *et al.*, 2005), no explicit localization of the negativities recorded on both partial errors and correct trials has been reported so far<sup>2</sup>. Obtaining similar localizations for these three activities would provide a strong argument in favor of the unicity of the phenomenon.

## 4.2 Materials and Methods

The data relative to the partial errors have been reported in a previous study for different purposes (see Burle *et al.*, 2008a where a detailed description of the experiment is available). The method will thus be briefly summarized, with emphasis on the aspects relevant for our current goals.

### 4.2.1 Subjects

Ten subjects aged from 20 to 31 years (mean : 25 years) volunteered for the experiment. All of them were right-handed and had normal or corrected-to-normal vision. According to the declaration of Helsinki, written informed consent before the start of the experiment was obtained from each subject.

### 4.2.2 Task, Recordings and EEG Data preprocessing

The subjects performed an Eriksen's flanker task (Eriksen et Eriksen, 1974). On each trial, three letters were presented to subjects who had to respond to the central one (target) while ignoring the others (distractors). They ran 20 experimental blocks of 128 trials each.

---

2. An indirect attempt was done by Vocat *et al.* (2008), however : although those authors did not explicitly attempt to localize the Ne-like on correct trials, they evaluated whether the solution found for errors could account for the activity on correct trials

Electroencephalographic activity (EEG) was recorded with 64 Ag/AgCl scalp electrodes (10-20 system positions, BIOSEMI Active-two electrodes, Amsterdam) and electromyographic activity (EMG) from the *flexor pollicis brevis* of each hand was recorded by paired surface Ag/AgCl electrodes. The sampling rate was 1024 Hz (filters : DC to 268 Hz, 3 dB/octave).

The vertical and horizontal EOG was recorded in order to correct eye movement artifacts by the statistical method of Gratton *et al.* (1983). All other artifacts were rejected after visual inspection of individual traces. The onset of the EMG activity was marked manually after visual inspection (for further details, see Burle *et al.*, 2008a).

The retained data were then downsampled to 256 Hz (with BrainAnalyzer, Munich), since the original sampling rate was too high to perform ICA.

#### 4.2.3 EEG Data analysis

The trials were sorted into three categories based on responses and on EMG patterns : pure-correct, error and partial error trials (see Burle *et al.*, 2002b, 2008a for more details). Partial errors trials are characterized by the presence of a small EMG burst on the incorrect response side preceding the EMG burst leading to the correct response. The EEG data were epoched, time-locked to the EMG activity that led to the overt response, namely the correct EMG burst for pure correct trials and partial errors, and the supraliminal incorrect EMG for errors.

Since previous reports have shown that the negativity on correct trials is much easier to observe after Laplacian computation, we also applied this transformation to the monopolar data, for the sake of comparison. The signal was interpolated with spherical spline interpolation, and hence the second derivatives in two dimensions of space were computed. We choose 3 for the degree of the spline since this value minimizes errors, and the interpolation was computed with a maximum of 15 degrees for the Legendre polynomial (Perrin *et al.*, 1989). We assumed a radius of 10 cm for the sphere representing the head, rather than the unrealistic default radius of 1 m assumed by BrainAnalyzer. With such a realistic radius, the most suitable unit is  $\mu V/cm^2$ .

#### 4.2.4 Blind Source Separation (BSS) : General Principle

ICA algorithm performs a “Blind Source Separation” (BSS) of the signal. Applied to EEG, BSS posits that the activities  $x_i(t)$  recorded on each sensors  $i$  (among  $I$ ) at time  $t$  can be decomposed as a sum of elementary components defined as the product of a topography ( $\alpha_{ij} \in \mathbb{R}^I$ , representing the contribution of the component to each of the  $i$  electrodes) and a time course  $s_j(t) \in \mathbb{R}^T$  (where

$I$  corresponds to the number of sensors,  $J$  the number of components with  $J \leq I$ , and  $T$  the number of time samples) :

$$x_i(t) = \sum_{j=1}^J \alpha_{ij} s_j(t) \quad (4.1)$$

Generally speaking, the goal of BSS is to recover both the  $\alpha_{ij}$  and the  $s_j(t)$  knowing only the realization across time of the signal  $x_i(t)$ . This problem can also be formalized in matrix terms and becomes :

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S} \quad (4.2)$$

where  $\mathbf{X} \in \mathbb{R}^{I \times T}$  is the observation matrix,  $\mathbf{A} \in \mathbb{R}^{I \times J}$  is the mixing matrix and  $\mathbf{S} \in \mathbb{R}^{J \times T}$  is the matrix of the components time course (*i.e.* the “sources” in the BSS terminology<sup>3</sup>).

Thus, BSS decomposes the input matrix  $\mathbf{X}$  into the product of the two matrices  $\mathbf{A}$  and  $\mathbf{S}$ . The decomposition of the  $\mathbf{X}$  matrix is not unique, however, and additional constraints are needed.

One widely used constraint is that the sources must be statistically maximally independent. Such an approach is usually called “Independent Component Analysis”. One way to quantify the independence between time series is to compute their mutual information (Bell et Sejnowski, 1995) and the **infomax** algorithm (used in the current study, see below) search for the  $\mathbf{S}$  matrix whose mutual information across components is minimal.

Applied to EEG, ICA aims at recovering *elementary* brain activities that are mixed at the sensors level because of volume conduction and diffusion effects, without any modeling of such conduction and diffusion effects (Jung *et al.*, 2001).

#### 4.2.5 ICA in the present study

Since, by demixing the observation matrix, ICA aims at recovering the *elementary* brain activities, we used ICA to address the unicity of the Ne and Ne-like observed on errors, partial errors and correct trials. Translated to ICA, the question of unicity becomes whether ICA decomposition can find a *single* component (although potentially with different temporal dynamic) accounting for the waves observed on the three categories of trials. In other words, can ICA *blindly* attribute

---

3. For the sake of clarity, we will consistently use the term “component” for the product of topography and time of course. We will use the word “generator” for the cortical activities reconstructed by source localization. Although, by applying ICA, one hopes that the recovered “components” will correspond to brain “generators”, it is important to keep the two concepts separated.



a single brain origin to the three waves? Addressing this question was performed in four steps, as detailed below.

**First step : construction of the matrices.** We first selected all the trials that were identified as correct, partial error or error trials. Partial errors were identified based on the EMG data. We removed trials contaminated by artifacts (see Burle *et al.*, 2008a for more details on artifact rejection). The selected trials were then segmented into epochs from -400 to +400 ms, centered on the EMG onset that triggered the overt response, either correct (correct and partial error trials) or incorrect (errors). Trials were then concatenated to form the  $\mathbf{X}$  matrix of size  $64 \times T$  (where 64 is the number of electrodes, and  $T$  is equal to `Number_of_trials`  $\times$  `Time_sample_per_trials`). One single  $\mathbf{X}$  matrix containing all the trials was built for each of the 10 subjects with the monopolar data. The individual matrices were of slightly different sizes since  $T$  depends on the number of trials included. Note that these data matrices are mainly composed of correct trials (76.8% on average) compared to errors and partial errors trials (5.2% and 18% on average, respectively).

**Second step : ICA on individual matrices.** ICA was applied to each of the  $\mathbf{X}$  matrices. ICA computations were performed with the `runica()` function as implemented in the EEGLAB software (Delorme et Makeig, 2004) which is derived from the Infomax ICA algorithm (Bell et Sejnowski, 1995). This returned 64 components per subject.

**Third step : searching for the component accounting for errors.** We searched for the component that could account for the Ne on errors (see Debener *et al.*, 2005). As already indicated, an component is defined as a topography and a time course. An component was thus selected if its averaged time-course *on errors* presented a clear phasic activity whose projection on the scalp was of negative polarity at fronto-central sites (see Fig. 4.2 below for an example). The selected component accounting for the Ne will be termed  $Ne\_IC_S$  for each subject  $S$ . Note that components were selected without any reference to correct trials.

**Fourth step : Averaging of the selected component for partial error and correct trials.**

Once the  $Ne\_IC_S$ , accounting for the Ne on errors, had been identified, we averaged the time course of  $Ne\_IC_S$  for partial errors and correct trials. Based on these averages, we sought whether the  $Ne\_IC_S$  could also account for the negativities observed on partial errors and on correct trials. To do so, we compared the component averages to both monopolar and Laplacian averages on the same trials.

### 4.2.6 Source Localization

To further probe similarities and differences between the negativities observed in the three categories of trials, we searched for the brain regions responsible for the genesis of those activities. Although source localization on errors has already been performed on several studies, to the best of our knowledge, the generator of the negativities on partial errors and on correct trials have never been localized so far. We thus sought for the generator of these three waves thanks to the standardized Low Resolution Electromagnetic Tomography (sLORETA) method, which implements a normalized form of the minimum norm constraint (see Pascual-Marqui, 2002 for technical details).

## 4.3 Results

### 4.3.1 Behavioral Data

Percentages of errors, partial errors and correct trials were respectively 5.2%, 18% and 76.8%, in average. The RTs associated with these three types of trials were significantly different ( $F(2,18) = 84.6, P < 0.001$ ), with the shortest RTs obtained for errors (342 msec), followed by pure-correct trials (382 msec) and partial errors (445 msec). The typical compatibility effects have been replicated, and are presented in more details in Burle *et al.* (2008a).

### 4.3.2 EEG Data

This section will be organized as follows : we will first present the monopolar data for comparison with the literature. We will then present the Laplacian data, since this transformation permits to reveal the negativity on correct trials. We will then present the ICA data, and finally the source localization results.

#### Monopolar data

Fig. 4.1A presents the monopolar grand averages obtained over FCz for errors (blue line), partial errors (green) and correct trials (red), time-locked to the relevant EMG activity onset (see above). The presented data replicates results already published, showing a clear negative wave for errors (Falkenstein *et al.*, 1991; Gehring *et al.*, 1993) and partial errors (Scheffers *et al.*, 1996; Vidal *et al.*, 2000) shortly after EMG onset. The amplitude of this wave was higher for errors ( $14.4 \mu V$ ) than for partial errors ( $7.05 \mu V$ ,  $F(1,9) = 13.26; p < 0.01$ ). As usually reported, no such phasic negativity is observable on correct trials (Fig. 4.1A, red line).

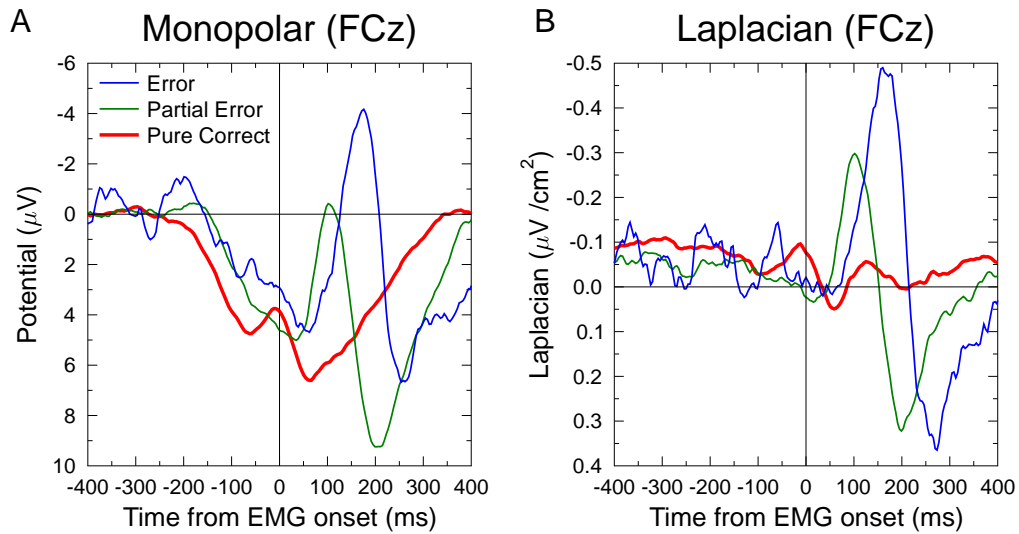


FIGURE 4.1 – Grand averages (FCz) of the three categories of trials (Error : blue, partial error : green and pure correct : red, for the monopolar data (panel A), and for the Laplacian transformed data (panel B). The zero of time indicates the relevant EMG onset : for correct trials, this corresponds to the EMG burst that triggered the correct mechanical response. For errors, it corresponds to the EMG burst that triggered the error. For partial errors, the relevant EMG is the small, subliminal, EMG burst occurring on the incorrect hand before the EMG burst triggering the correct response. On the monopolar data (panel A), one can clearly see the Ne for errors and partial errors, but no negativity is visible on correct trials just after EMG onset. As already reported, a large positivity is observable, instead. On the contrary, after Laplacian transform, a clear negativity, although of smaller amplitude, appears also for correct trials. The large positivity, likely generated by remote sources, has disappeared, hence allowing to reveal the small negative wave, much more focal, present on correct trials.

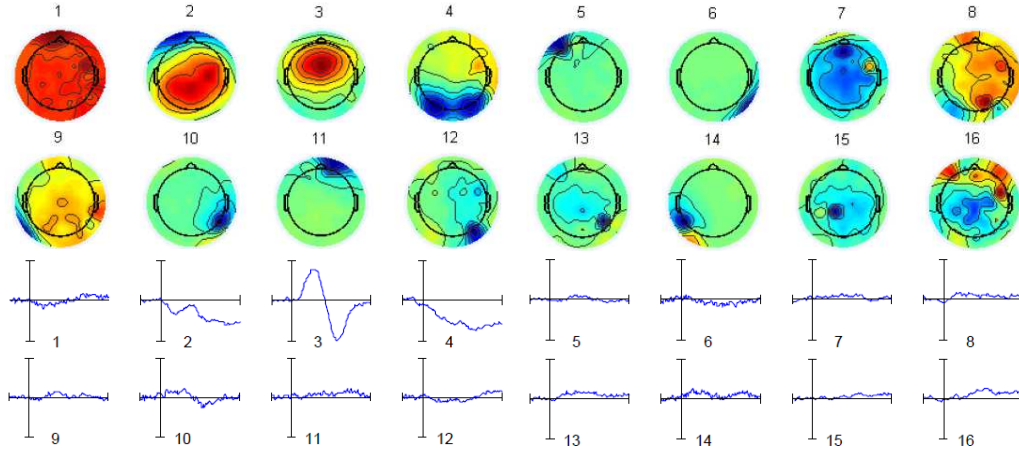


FIGURE 4.2 – **Example of decomposition and component selection for subject 01.** This figure presents the 16 first components (among 64). The first two row present the topographies of the components, and the last two rows present the corresponding averaged time courses for the same components for error trials, time locked to incorrect EMG onset. In this example, the third component clearly presents a fronto-central distribution and a clear increase in activity just after the incorrect EMG. This component was thus selected for this particular subject. The same procedure was applied to all the subjects.

### Laplacian data

The Laplacian data provide a different picture : although negative waves are still clearly observable on errors and partial errors (Fig. 4.1B, lines blue and green, respectively) over FCz, a small negative activity now becomes visible on correct trials over the same electrode (red line on Fig. 4.1B), with a latency comparable to that observed for errors and partial errors. The amplitude of those negativities significantly differ across trial types ( $F(2, 18) = 34.9, P < 0.001$ ). A latency effect of the peak also shows up ( $F(2, 18) = 24.5, P < 0.001$ ), with a peak occurring earlier for partial error (mean latency : 109 ms) than for correct trials (mean latency : 123 ms) and errors (mean latency : 162 ms). The latency difference between partial and full errors replicates results already reported by Carbonnell et Falkenstein (2006) and extend them, by showing that the latency for correct trials lies in between partial and overt errors.

### 4.3.3 ICA Data

Fig. 4.2 illustrates how the component  $Ne\_IC_{01}$  was selected for subject 01. This figure presents the 16 first components, with their topography (top) and their temporal dynamics on error trials

(bottom, the vertical bar indicates EMG onset). As one can easily see, the component n° 3 presents a topography and a temporal dynamic compatible with the Ne, with a large activity occurring shortly after EMG onset. This component was thus chosen as representing the Ne on errors, and will be called  $Ne\_IC_{01}$ , where “01” stands for subject 01. The same procedure was applied to all the subjects, allowing to identify the  $Ne\_IC_S$ ,  $S \in \{1, 10\}$ . Note that, during this selection stage, the time course of the components for correct and partial error trials was unknown.

Based on those criteria, we could find a component  $Ne\_IC_S$  whose topography and time course could account for the Ne on errors (First step described above) for each subject. Fig. 4.3 shows the topography of all the selected  $Ne\_IC_S$  for each subject  $S$  and table 4.1 gives the rank of the selected  $Ne\_IC_S$  for each subject  $S$ . It is to be noted that the  $Ne\_IC_S$  are always among the first components, for all the subjects (lowest rank = 9). As components are ranked according to their energy, this indicates that they account for a large part of the original signal.

Once each of  $Ne\_IC_S$  had been identified, we averaged, for each subject  $S$ , the time courses of these  $Ne\_IC_S$  for correct trials and partial errors. The grand average is presented in Fig. 4.4A. One can clearly see a negative activity for both partial errors and, more importantly, for correct trials, albeit of smaller amplitude<sup>4</sup>. Statistical analysis revealed that the amplitudes of the peaks depend on trial types ( $F(2, 18) = 24.0, P < 0.001$ ). A latency peak effect also showed-up ( $F(2, 18) = 34.5, P < 0.001$ , mean latencies : 116, 111 and 162 ms for correct, partial errors and errors, respectively). For the sake of comparison, Fig. 4.4B also presents the Laplacian data plotted with the same baseline and focus. The similar statistical results, along with the comparison between panels A and B indicate a close similarity between the Laplacian data and the component isolated by ICA. We will come back on this similarity in the discussion.

#### 4.3.4 Source Localization

##### Localization of the selected component

Assuming that each  $Ne\_IC_S$  reflects a single neural source, one can attempt to localize the generator of the component with a single equivalent dipole. This was done for each  $Ne\_IC_S$  (*i.e.* for each subject) with the “**dipfit**” module of EEGLAB. Fig. 4.5 shows the equivalent dipole located at the median position of the individual dipoles. These median positions ( $x = 0.6$ ,  $y = 4$  and  $z = 36.2$ , Talairach coordinates) clearly point to a source in the Rostral Cingulate Zone

4. Since the data matrix was centered on the response-related EMG, the available pre-EMG activity for partial error is shorter than for the other trials. This is why only 100 ms before EMG onset is presented

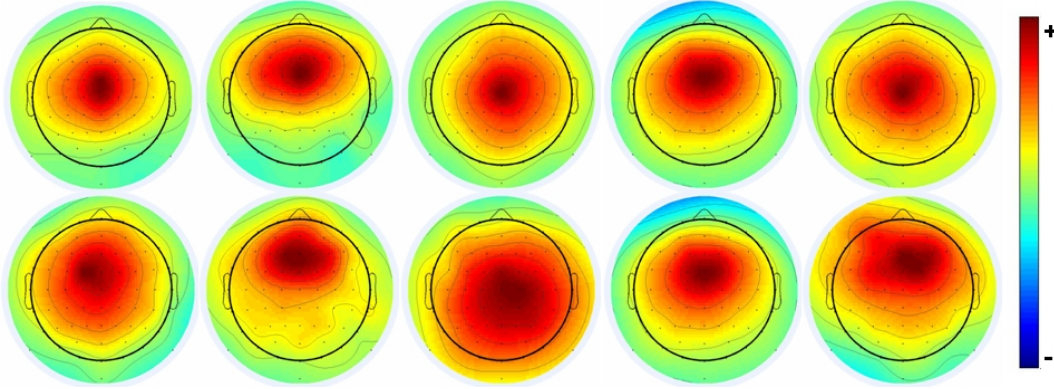


FIGURE 4.3 – **Topographies of the  $Ne\_IC_S$  for each of the 10 subjects.** All of them present a fronto-central topography. Note that, although all the maps present a positive polarity, this polarity is somehow arbitrary. Indeed, the projected signal on the sensors space is the product of the topography and the time-course. Hence inverting the polarity of both the topography and the time course produces exactly the same results. For the sake of simplicity, all the topographies have thus been plotted as positive, and the time course as negative in the period of interest. Importantly, for all the  $Ne\_IC_S$ , the projected activity in the sensors space was always of negative polarity at the time of the Ne.

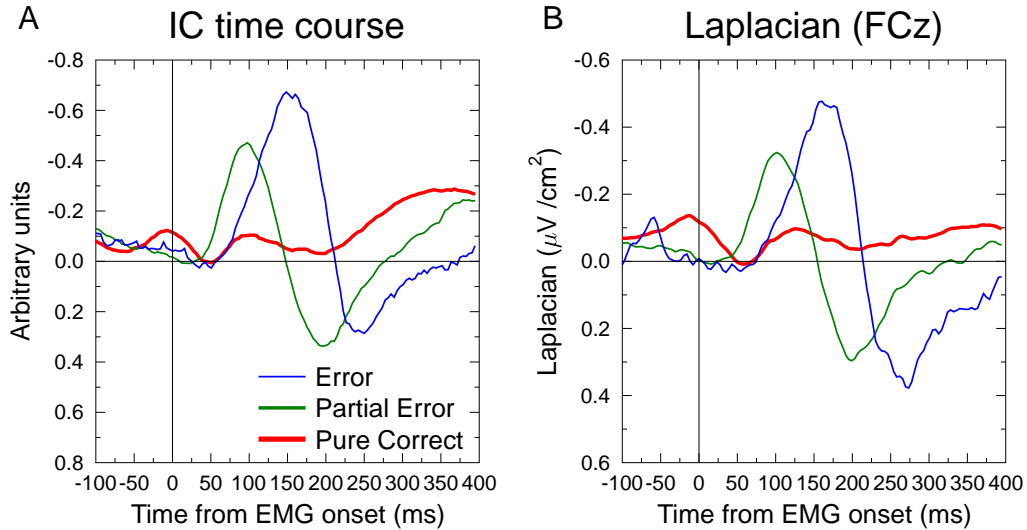


FIGURE 4.4 – **Grand average of the  $Ne\_IC_S$  time course (panel A) and of the Laplacian transformed data (panel B, electrode FCz), for the three categories of trials.** The zero of time corresponds to EMG onset (see Fig. 4.1 for details). The ICA data present a clear negative activity just after EMG onset, for both the partial errors and the correct trials.

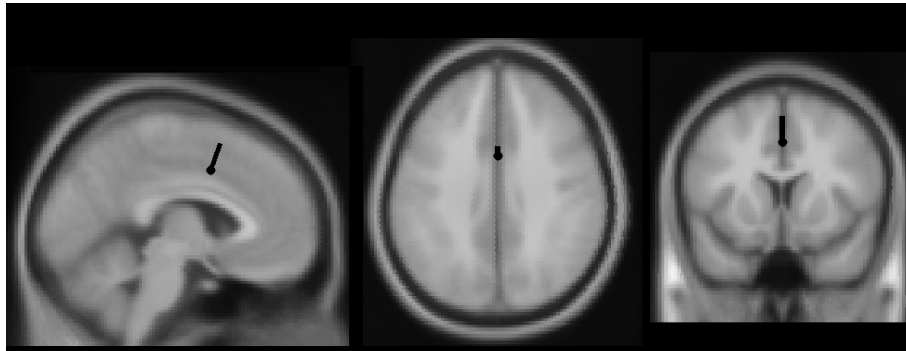


FIGURE 4.5 – **Source localization of the  $Ne_{ICs}$ .** A localization of all the  $Ne_{ICs}$  was performed with *dipfit* plugins of EEGLAB. The median position of all those dipoles coordinates and orientation was taken, and the equivalent median dipole is plotted in the MNI averaged brain. The equivalent dipole accounting for the  $Ne_{ICs}$  is located within the RCZ.

(Ridderinkhof *et al.*, 2004a) which is very likely the generator of the Ne (Dehaene *et al.*, 1994; Van Veen et Carter, 2002; Debener *et al.*, 2005; Herrmann *et al.*, 2004).

#### Distributed sources Localization on the original data

The ICA data suggest that the negativities observed on errors, partial errors and correct trials are of same nature. To further establish this point, we performed a source localization of the EEG data in the time range of those negativities. The Ne on errors has already been localized, both with single dipole modeling (Dehaene *et al.*, 1994; Van Veen et Carter, 2002) and with distributed sources (Herrmann *et al.*, 2004). There is a clear agreement for a source in the RCZ, although the precise structure might still be debated.

In the present study, Source localization was performed on the grand average by using the standardized Low Resolution Electromagnetic Tomography (sLORETA) method (Pascual-Marqui, 2002), which implements a particular normalization of the minimum norm algorithm. Sources were computed at 110 ms after EMG onset. Fig. 4.6A confirms that the source of the Ne on errors is within the RCZ, in the ACC and/or in the SMA. Importantly very similar sources were found for the two other types of trials. Indeed, for both the partial errors and the correct trials, sLORETA solutions are also clearly localized in the RCZ region.

#### 4.3.5 Lateralization of the $N_C$ on correct trials ?

Yordanova *et al.* (2004) observed that on correct trials the Laplacian topography had a fronto central dominance with “[...] an additional tendency for a greater involvement of contralateral

Subject	Rank	R. V.	Tal. Coord.			MNI Coord.		
			x	y	z	x	y	z
01	3	4.73	0.7	5.7	36.7	46.4	66.0	57.1
02	4	2.67	0.6	-15.6	20.1	46.3	55.4	47.5
03	5	5.51	0.9	25.3	35.7	46.5	76.1	57.0
04	1	2.56	0.6	-17.4	2.5	46.3	55.0	37.9
05	2	1.99	0.6	-13.9	37.6	46.3	55.9	57.1
06	4	2.95	-15.1	5.9	36.7	38.4	66.1	57.1
07	9	4.53	0.9	23.9	18.1	46.5	75.7	47.5
08	4	5.25	0.7	2.2	1.5	46.4	65.1	37.9
09	2	3.66	0.7	5.7	36.7	46.4	66.0	57.1
10	4	5.40	0.7	-13.9	37.6	46.3	55.9	57.1
Median	-	-	0.6	4.0	36.2	46.3	65.6	57.1

TABLE 4.1 – **Summary of the equivalent dipole fitting for all the subjects.** For each subject, this table indicates the rank of the  $Ne\_IC_S$ , the residual variance (R. V.), and the dipole coordinates, in both Talairach and MNI coordinates.

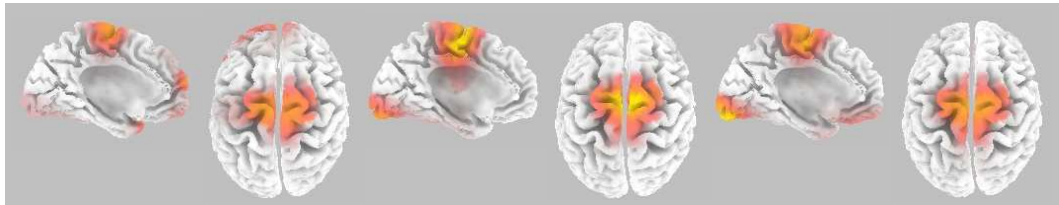


FIGURE 4.6 – **Source localization results obtained with sLORETA for errors (left), partial errors (middle) and correct (right) trials.** For the three trial types, a clear source within the RCZ was obtained.



regions [...]” (p. 594), whereas it was central on errors. This led them to conclude that the generators of the negativities in correct and error trials differ, at least partially. The finding that a single component could account for the negativities in the three categories of trials clearly speaks against different generators. Here, we explored another hypothesis, namely that the observed lateralization is due to another, lateralized, generator. If one assumes that the strength of this second generator is independent of correctness, it will largely impact the observed topography on correct trials (where the central activity is small), whereas its effect will be limited when the central activity is large (*i.e.* for errors). In order to evaluate this hypothesis we measured the activity over the FC1/FC2 electrodes (depending on response side). As a matter of fact, a clear negativity was observed over these electrodes for the three categories of trials (Fig. 4.7). Its surface (between 90 and 110 ms, around the peaks) did not vary as a function of correctness ( $-0.195$ ,  $-0.201$  and  $-0.144 \mu V/cm^2$  for correct, partial errors and errors, respectively,  $F(2, 18) = 0.93$ ,  $P = 0.41$ ), however. Furthermore, although the medio-central negativities present a latency shift (see above), the latency of the lateral negativities did not (103 ms, 98 ms and 104 for correct, partial errors and errors, respectively,  $F(2, 18) = 0.49$ ,  $P = 0.62$ ). This analysis confirms the presence of a negativity contralateral to the produced EMG activity for all three types of trials. Its amplitude being the same for the three categories, the lateralization is more visible for correct trials. This impression is further amplified by latency effects (see Fig. 4.8). Indeed, the FC2 activity peaking earlier than the Ne on errors, the topography at the peak of the Ne is less influenced by the lateralized activity. On the contrary, partial errors, and even more, correct trials present a greater lateralization since the time-course of the medial component is similar to the lateralized one.

## 4.4 Discussion

Since its discovery by Falkenstein and colleagues (Falkenstein *et al.*, 1991), the Ne has attracted a lot of interest in the cognitive control literature. Yet, its precise functional role is still unclear (Burle *et al.*, 2008a). Among open issues, the question as to whether the “Ne-like” observed on correct trials (Vidal *et al.*, 2000, 2003a; Mathalon *et al.*, 2002) is of the same nature as the Ne observed on errors and partial errors, or whether it reflects a completely different activity, is critical. We addressed this point by using the property of ICA to *blindly* identify the *elementary components* of the EEG signal. We reasoned that if the Ne on errors and on correct trials reflect the same brain activity modulated in amplitude, it should be captured in a single component by ICA. On the contrary, these negativities should be scattered into different components if they reflect completely

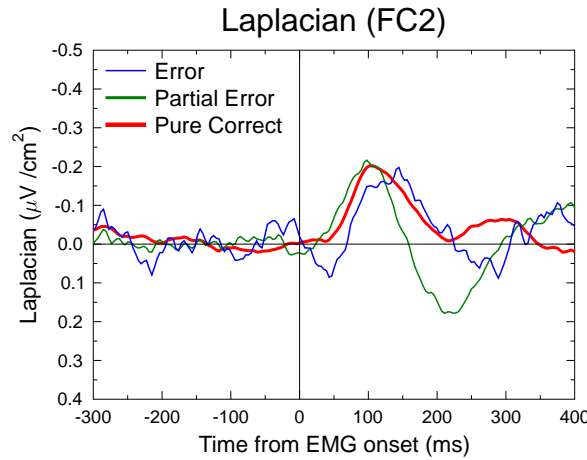


FIGURE 4.7 – Grand average of the fronto-lateral negativity observed over FC1/FC2 (depending on response side) after Laplacian computation, for errors (blue), partial errors (green) and correct (red) trials. This lateral activity is present for the three categories of trials, and neither its amplitude nor its latency differ across the three trial types. This activity is likely a follow-up of the negativity observed above the primary motor cortices contralateral to the response, just before EMG onset (see Burle *et al.*, 2008a, Burle *et al.*, 2004b for an overview).

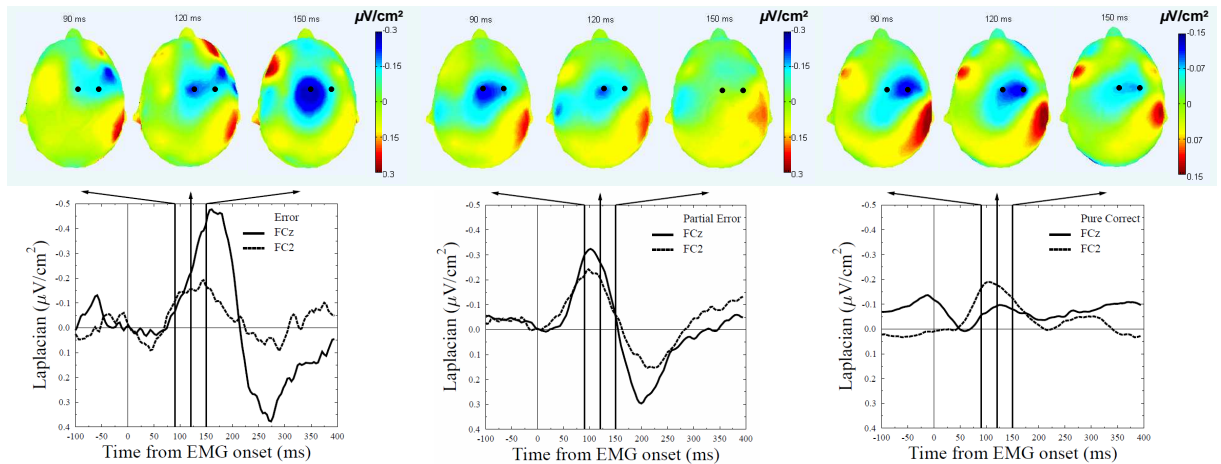


FIGURE 4.8 – Topographies and time courses of the medial and lateral activity, for the three types of trials (error, partial errors and correct trials, from left to right, respectively). The enhanced lateralization on correct comes from two interacting factors : first, on correct, the amplitude of the lateralized activity being closer (actually even larger) than the medial one, it largely contributes to the topography. The larger amplitudes for the medial activity on partial errors and errors, while the lateralized one does not change, reduces the impression of lateralization, although the lateralized component is of same amplitude. The second factors is timing. Indeed, the latencies of the medial activity differs across trial types, but it is not the case for the lateralized one. The peaks of the medial and lateralized activity are very similar for correct, but less so for errors, which further emphasizes the impact of the lateralized activity on correct trials.

different brain processes.

For all subjects, ICA could isolate a single component, localized within the RCZ (Fig. 4.5), accounting for the negativities in the three categories of trials (correct, partial errors and errors). In other words, ICA *blindly* attributed a single source to the negativities recorded for those three categories of trials.

Besides the ICA argument, the source localization results also point toward a unique brain activity modulated in amplitude. As already reported (Dehaene *et al.*, 1994; Van Veen et Carter, 2002; Herrmann *et al.*, 2004), the generator of the Ne on errors was located within the RCZ. The present data extend these results by locating, for the first time, the  $N_C$  within the RCZ, for both partial error and correct trials. Although a common brain source localization cannot be considered as a definitive argument for the unicity of the underlying process, this provides strong support for such a claim. Furthermore, the source of the  $N_C$  being located within the RCZ, known to be largely involved in cognitive control (Ridderinkhof *et al.*, 2004a), further supports the interpretation of the  $N_C$  in terms of cognitive control, even on correct trials, in agreement with previous data showing its sensitivity to performance (Luu *et al.*, 2000b; Allain *et al.*, 2004c). For example, Allain *et al.* (2004c), capitalizing on previous results (Ridderinkhof *et al.*, 2003), have shown that the  $N_C$  amplitude was reduced on correct trials preceding an error, hence foreshadowing the forthcoming error.

Yordanova *et al.* (2004) disputed the unicity in reporting that the  $N_C$  on correct trials was lateralized contralaterally to the produced response while the Ne on error trials was not. The present data confirm the lateralization, but also show that such a lateralization is not specific to correct trials, as it is also present on partial errors and error trials. The amplitude of the lateralized component being the same for the three type of trials (Fig. 4.7), scaling effects make this lateralization appear much stronger on correct trials. Differences in time courses further exaggerate this impression. The differential modulations, both in latency and in amplitude, of the medial and lateral components across trial types clearly indicates that they reflect different brain generators, hence confirming that the  $N_C$  is medial, as the Ne on errors, and not lateralized.

Altogether, our results clearly point to a single process whose amplitude is modulated by performance. This has important consequences for our understanding of the functional role of the RCZ, and for cognitive control in general. Indeed, instead of being specific to error (Coles *et al.*, 2001), or to conflict trials (Yeung *et al.*, 2004), it appears that the RCZ is recruited by every response activation, but that the strength of this recruitment depends on the correctness of response activation and/or the time needed to correct and erroneous activation (Burle *et al.*,

2008a). This novel view adds very strong constraints on formal models of cognitive control.

Last, our results call for a more methodological comment. The results obtained with ICA and Laplacian transform are remarkably similar. Although both techniques aim at separating the contribution of the underlying sources, they are, however, extremely different from a mathematical point of view. Indeed, ICA decomposition is mainly based on the *temporal* statistical dependence between signals recorded across sensors, while Laplacian is the second *spatial* derivative of the recorded signal. The high convergence between these two different methods cross-validates their results and strengthens the above results. It also exemplifies the danger of deriving conclusions on monopolar, mixed data, and strongly speaks for the use of deblurring methods in EEG.

## Chapitre 5

# Sélection de la réponse

### 5.1 Introduction

Dans la vie quotidienne, nous sommes très souvent amenés à prendre des décisions. Certaines nécessitent une longue réflexion tandis que d'autres doivent être prises très rapidement. Les mécanismes neuronaux impliqués dans ces deux situations sont certainement très différents. Nous nous intéressons ici à la prise de décision rapide à laquelle nous faisons appel lorsque, par exemple, nous sommes au volant d'un véhicule. En laboratoire, cette prise de décision rapide est souvent étudiée dans des tâches de Temps de Réaction (TR) à plusieurs éventualités. La pression temporelle à laquelle sont soumis les sujets, les incite à prendre leur décision le plus vite possible. Exécuter correctement la bonne réponse dans les tâches de TR de choix implique que le stimulus a été identifié, que les règles d'associations stimulus-réponse ont été utilisées pour sélectionner la réponse appropriée, et enfin que le programme moteur a été exécuté. Comment le système nerveux parvient-il à faire son choix de réponse lorsque deux réponses différentes peuvent être sélectionnées ? Ce problème intéresse les neuroscientifiques depuis longtemps. De nombreuses hypothèses ont été proposées sur le déroulement des étapes de traitement de l'information. Depuis quelques années, certaines de ces hypothèses sont implémentées dans des modèles neuromimétiques. L'étude des modèles présentée dans le chapitre 2 fait ressortir des modes de sélection de la réponse très différents. Notamment, nous pouvons distinguer ces modèles sur la base de la présence ou non d'une étape d'association Stimulus-Réponse (S-R).

Pour une classe de modèle, la décision sur la réponse est intrinsèquement liée à la décision sur le stimulus (par exemple la lettre H ou S pour la tâche d'Eriksen, voir figure 2.6 page 51). Dans

cette classe de modèle (Usher et McClelland, 2001; Bogacz *et al.*, 2006; Ratcliff *et al.*, 1999), le choix de la réponse s’effectue directement au niveau des structures qui implémentent l’exécution de la réponse et les unités de réponses accumulent, au cours du temps, des preuves en faveur de l’un ou l’autre des stimuli. Lorsque le niveau d’activité d’une des unités de réponse atteint un certain seuil, la réponse est donnée. Dans ce type de modèle, aucune erreur d’association S-R ne peut être commise, seules des erreurs de catégorisation perceptives dues au bruit ou à une interférence liée au stimulus, peuvent apparaître. Ce type de prise de décision est implémenté dans la partie “tâche” utilisée par le modèle du conflit. C’est également le cas pour le modèle *WEAVER++* dans lequel la réponse motrice est automatiquement programmée dès l’identification de la cible quelle que soit la modalité de réponse (manuelle ou de production orale, voir figures 2.11 et 2.14 pages 61 et 66).

Une autre classe de modèles considère au contraire que l’interface entre stimuli et réponse se fait par un processus spécifique d’association. Par exemple, le mode de sélection de la réponse du modèle *ERN-RL* propose une distinction entre le choix de la cible (couche catégorielle) et le choix de la réponse motrice (couche de réponse). Les règles d’association S-R sont inscrites dans les valeurs des poids qui relient les unités de ces deux couches (voir figure 2.16 page 70). Les deux aspects de la règle (identification de la cible et association S-R) sont dissociés. Bien que non décrit en ces termes, l’analyse des équations montre que ce modèle présume que la transmission de l’information est discontinue, et que l’enchaînement des étapes se fait de manière séquentielle. Puisque l’architecture de ce modèle suppose que l’étape de sélection de la cible précède l’étape d’association S-R, seule l’information sur la cible est “traduite” en réponse motrice. Les deux étapes sont également dissociées dans le modèle des GDB de (Frank, 2005). Chaque stimulus est codé par les neurones du striatum. Les auteurs proposent que la sélection de la réponse se fasse au niveau des connexions entre le striatum et le globus pallidus interne, et entre le striatum et le globus pallidus externe (voir figure 2.17 page 73).

La sélection de la réponse à surtout été étudiée en imagerie cérébrale métabolique. Il a été proposé que les règles d’associations S-R soient implémentées au niveau du cortex pariétal supérieur et/ou du cortex dorsal préfrontal (Jiang et Kanwisher, 2003; Schumacher et D’Esposito, 2002; Heekeren *et al.*, 2004). Malheureusement, les aspects temporels des activations sont, du fait de la technique, mal connus. Or connaître la dynamique temporelle des activations des aires impliquées dans ces processus de prise de décision est nécessaire à la compréhension de leur fonctionnement. Récemment, l’utilisation du Laplacien de surface a permis d’étudier plus finement les activités cérébrales prenant place juste avant la production de la réponse, ouvrant ainsi la possibilité d’étudier plus directement les processus impliqués dans la sélection et l’exécution des

réponses. Notamment, Vidal *et al.* (2003b) ont montré qu'il était possible de dissocier, spatialement et temporellement, les activités produites par les cortex moteurs primaires (M1) et l'aire motrice supplémentaire (AMS). L'application du Laplacien de surface sur les données synchronisées sur la réponse a révélé, au dessus des cortex moteurs primaires, une négativité et une positivité respectivement sur l'hémisphère controlatéral et ipsilatéral à la réponse (voir figure 1.12 page 28 et Taniguchi *et al.* (2001); Tandonnet *et al.* (2003); Vidal *et al.* (2003b)). Ces deux activités démarrent une centaine de millisecondes avant le début de l'EMG. La négativité sur M1 controlatéral a été interprétée comme reflétant une "activation" du cortex moteur permettant l'envoi de la commande motrice. La positivité sur M1 ipsilatéral a été interprétée comme une "inhibition" du cortex moteur impliqué dans l'exécution de la réponse incorrecte, et dont le rôle fonctionnel pourrait être de prévenir les erreurs (Vidal *et al.*, 2003b; Hasbroucq *et al.*, 2000; Burle *et al.*, 2002a, 2004b).

Précédant ces activités issues des cortex moteurs primaires, Vidal *et al.* (2003b) ont mis en évidence une activité de polarité négative qui se développe aux alentours de 100 ms avant le début de l'EMG et dont le maximum est mesuré aux alentours de -40 ms. Cette négativité, que nous nommerons N-40, est enregistrée en regard de l'électrode FCz<sup>1</sup> et précède les activités observées au dessus des cortex M1 (voir figure 1.12 page 28). Elle reflète probablement l'activité de l'AMS. Carbonnell *et al.* (2004) se sont intéressés au rôle fonctionnel de cette N-40, et se sont demandés si cette onde pouvait être impliquée dans les processus de sélection de la réponse. Ils ont montré que l'amplitude de la N-40 était réduite lorsque les sujets savaient à l'avance quelle réponse ils allaient devoir effectuer. Ce résultat suggère que cette diminution est due au transfert d'une partie de la sélection et de la programmation de la réponse pendant la période préparatoire. D'autre part, la N-40 n'a pas été observée dans une tâche de TR simple dans laquelle le sujet doit appuyer sur un bouton le plus vite possible dès l'apparition d'un signal de réponse (Burle *et al.*, 2004b). Dans cette tâche, il n'y a pas de choix à faire puisque le sujet sait à l'avance la réponse qu'il va devoir exécuter.

Le patron d'activation-inhibition et la N-40 semblent deux bons candidats électrophysiologiques pour étudier les processus de prise décision. En effet, si, comme le posent les modèles sans étape d'association S-R, le choix de la réponse s'effectue directement au sein des cortex M1, alors les cortex M1 sont le lieu de la compétition entre les réponses. La forme du patron d'activation-inhibition devrait être différente dans les essais corrects compatibles et incompatibles. En effet, du fait de l'architecture du modèle, l'activation de la réponse incorrecte par les distracteurs devrait ralentir l'accumulation des "preuves" par l'unité de réponse correcte, nous devrions observer un patron

---

1. électrode fronto-centrale sur la ligne médiane

d'activation-inhibition qui démarre plus tôt par rapport à la réponse dans les essais incompatibles que dans les essais compatibles. Si au contraire, une étape d'association S-R (et donc de sélection de la réponse) existe en amont des cortex moteurs, alors le patron d'activation-inhibition ne devrait pas être affecté par la compatibilité du stimulus. On devrait par contre observer une activité qui précède l'activité des M1, et qui soit sensible à la règle d'association S-R. La N-40, du fait de sa latence et de sa localisation anatomo-fonctionnelle pourrait être un candidat pour cette étape de sélection. Si la N-40 est bien impliquée dans les processus de sélection de la réponse, alors son rôle pourrait être d'appliquer les règles d'association S-R. Dans les essais incompatibles, la compétition entre les deux réponses est grande, cette étape d'association S-R devrait être davantage sollicitée. Si cette hypothèse est valide, nous devrions observer une augmentation de l'amplitude de la N-40 dans la condition incompatible par rapport à la condition compatible.

Nous avons donc cherché à préciser les processus impliqués dans le choix entre deux réponses en manipulant la compatibilité du stimulus et en observant l'influence de cette manipulation sur la N-40 et le patron d'activation-inhibition. Nous avons étudié cette question dans deux tâches manipulant la compatibilité, mais différant à la fois par la nature des stimuli ainsi que par l'interférence, afin d'évaluer la robustesse des résultats aux conditions particulières de la tâche.

## 5.2 Expérience 1

### 5.2.1 Introduction

Dans cette première étude, nous avons utilisé une tâche de Simon dans sa version dite “accessoire”. Les stimuli étaient composés d'une lumière présentée centralement et d'un son présenté latéralement dans une des deux oreilles. Les sujets devaient répondre en appuyant sur un bouton avec le pouce droit et le pouce gauche, ce qui permet de dissocier le cortex moteur impliqué dans la bonne réponse et le cortex moteur impliqué dans la réponse alternative. Les résultats présentés ci-dessous concernent uniquement l'analyse des activités liées à la réponse.

### 5.2.2 Matériel - Méthode

#### Les sujets

Seize sujets (âgés de 21 à 34 ans, moyenne : 26.5 ans) ont participé à l'expérience. Ils avaient une vision normale, ou normale après correction.



### Tâche et procédure

Les sujets, confortablement installés dans un fauteuil placé dans une cage de Faraday, effectuaient une tâche de temps de réaction de choix à deux éventualités.

Deux stimuli étaient présentés simultanément : un stimulus visuel et un stimulus auditif. Le stimulus visuel était une lumière de couleur rouge ou verte qui apparaissait au centre d'un panneau noir dans le champ de vision central du sujet. Le stimulus auditif était présenté soit dans l'oreille gauche, soit dans l'oreille droite du sujet. Pour présenter les sons, nous utilisons un dispositif amagnétique pour chaque oreille. Les sons entendus dans une oreille ne pouvaient pas être entendus par l'autre.

La consigne était de répondre le plus vite possible en fonction de la couleur du stimulus visuel en appuyant sur le bouton placé dans la main droite ou dans la main gauche.

Le nombre d'essais par bloc était de 80. Chaque sujet effectuait 2 blocs d'entraînement et 20 blocs expérimentaux. Les essais étaient séparés d'un intervalle inter-essais constant de 1500 ms.

### Recueil des données

Les données comportementales recueillies étaient le temps de réaction et le taux d'erreur. Les activités l'électroencéphalographiques et électromyographiques ont été enregistrées grâce à un système d'acquisition BIOSEMI utilisant des électrodes d'enregistrement pré-amplifiées (filtrage : DC à 268 Hz, 3dB/octave). L'activité électromyographique (EMG) des deux mains a été enregistrée grâce à deux électrodes (Ag/AgCl) collées sur la peau de l'éminence Thénar au-dessus du *flexor policis brevis* impliqué dans la flexion du pouce (voir figure 5.2). L'électroencéphalogramme (EEG) a été enregistré grâce à 64 électrodes (Ag/AgCl) disposées sur le cuir chevelu (système 10-20). L'électrode de référence était placée sur la mastoïde gauche. L'EEG conventionnel possède une très mauvaise résolution spatiale (9-10 cm). Il est néanmoins possible d'améliorer considérablement cette résolution jusqu'à 2-3 cm grâce au calcul du Laplacien de surface (Babiloni *et al.*, 2001) qui fournit une bonne approximation du corticogramme, c'est à dire de l'activité qui serait enregistrée à la surface du cortex. Le Laplacien est obtenu pour l'ensemble des électrodes. Nous nous focalisons néanmoins sur les électrodes qui se trouvent en regard du CCA et/ou de l'AMS et des cortex moteurs, c'est à dire FCz, C3 et C4. Nous nous sommes également intéressés aux activités au dessus des aires prémotrices (électrodes FC1/FC2). Les mouvements oculaires ont été enregistrés grâce à des montages bipolaires entre deux électrodes placées au-dessus et en dessous de l'œil gauche pour les mouvements verticaux (EOGV) et entre les deux canthi externes pour les mouvements

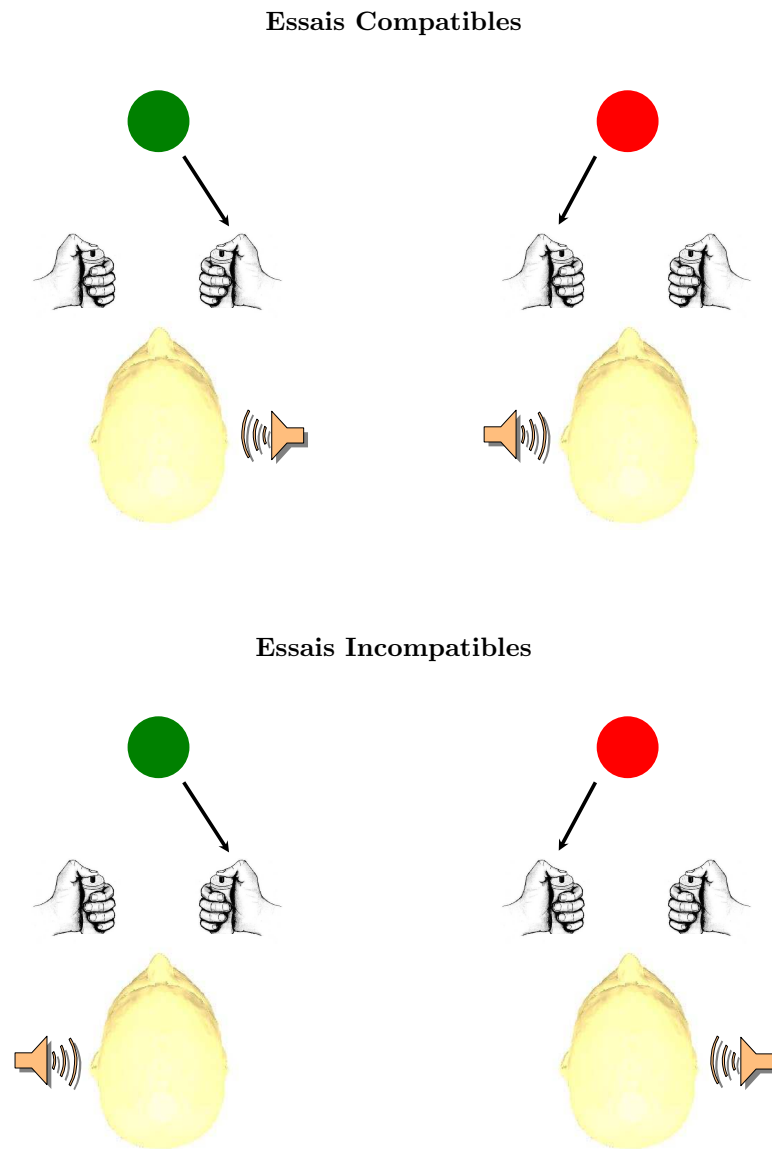


FIGURE 5.1 – **Illustration de la tâche de Simon dans sa version dite “accessoire”**. Le sujet doit identifier la couleur du stimulus et, comme dans cet exemple, répondre le plus vite possible avec le bouton de droite si la couleur est verte et avec le bouton de gauche si la couleur est rouge. La localisation du son n’est pas pertinente dans cette tâche. Les essais compatibles sont ceux pour lesquels le son est présenté du côté de la réponse correcte (en haut), et les essais incompatibles sont ceux pour lesquels le son est présenté du côté de la réponse incorrecte (en bas).

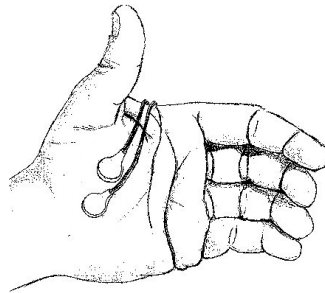


FIGURE 5.2 – Placement des électrodes EMG sur la peau de l'éminence Thénar au-dessus du *flexor pollicis brevis*

horizontaux (EOGH).

### Traitement des données

Les tracés électrophysiologiques ont été filtrés entre 0.3 et 100 Hz pour l'analyse de l'EEG, et à 10 Hz (passe haut) pour l'analyse de l'EMG.

**Traitements EMG** Les activations EMG ont été inspectées visuellement et chaque bouffée a été marquée manuellement. De cette façon, nous avons pu classer les essais en trois catégories. Tout d'abord, nous avons séparé les essais corrects et les erreurs. Puis, nous avons distingué deux catégories dans les essais corrects : les essais “pur-corrects” présentant uniquement une bouffée EMG correcte avant la réponse, et les “ébauches d'erreur” présentant une petite activation EMG du muscle impliqué dans la réponse incorrecte (EMG incorrect) suivie d'une activation EMG du côté de la réponse correcte (EMG correct) déclenchant la réponse.

**Prétraitements EEG** Les artéfacts présents dans les tracés ont été rejetés par inspection visuelle de tous les tracés de toutes les électrodes sur tous les essais. Les artéfacts induits par les mouvements oculaires ont été soustraits par une méthode statistique (Gratton *et al.*, 1983). Toutes ces identifications, bien que fastidieuses, sont nécessaires car les algorithmes automatiques chargés du rejet d'artéfacts et de la sélection des mouvements oculaires ne sont pas satisfaisants.

**Traitements EEG** Les tracés EEG ont été segmentés en prenant comme zéro des temps le début de l'EMG, puis triés en fonction de la compatibilité du stimulus et de la nature de la réponse (pure-correcte, correcte avec ébauche d'erreur ou erreur). Nous nous concentrerons ici sur les essais pur-corrects obtenus pour les essais compatibles et incompatibles, que nous analyserons

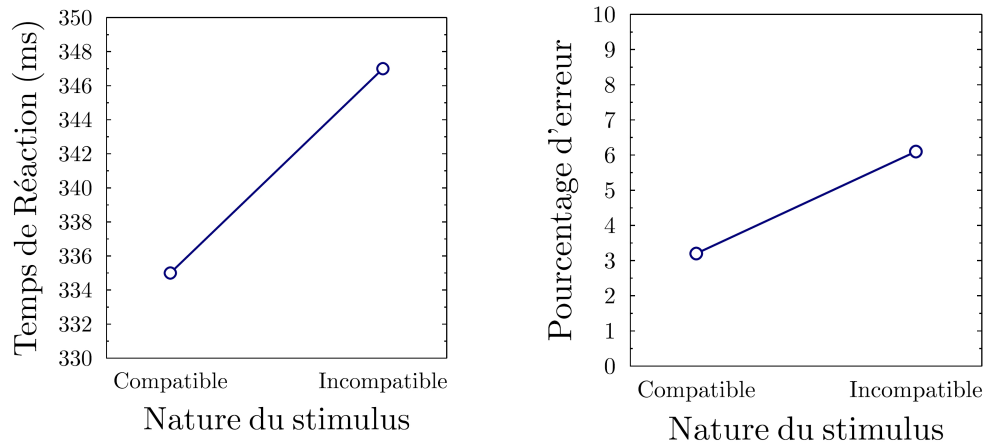


FIGURE 5.3 – **Expérience 1 : Résultats Comportementaux.** Temps de réaction moyens et pourcentage d’erreur moyens chez les sujets en fonction de la compatibilité du stimulus.

séparément.

Le Laplacien de surface a été calculé sur la base de ces segments moyennés sur l’ensemble des électrodes. La méthode utilisée est celle développée par Perrin *et al.* (1989) et implémentée dans BrainAnalyser<sup>©</sup> (Munich, Allemagne). Dans un premier temps une interpolation spline sphérique est effectuée (ordre des splines : 3 ; degré maximum du polynôme de Legendre : 15), puis la double dérivation dans les deux dimensions du plan est calculée. L’activité électroencéphalographique après calcul du Laplacien de surface est exprimée en  $\mu V/cm^2$ .

### 5.2.3 Résultats Comportementaux

Les pourcentages d’essais pur-corrects, corrects avec ébauche d’erreur et erreurs sont respectivement de 83.24%, 12.60% et de 4.16%.

#### Temps de Réactions

Nous nous sommes intéressés aux variations de TR des essais corrects en fonction de la compatibilité du stimulus. Parmi les essais corrects, les TR des essais compatibles sont significativement plus courts que les TR des essais incompatibles (335 versus 347 ms,  $F_{(1,15)} = 37.7, p < 0.001$ , voir figure 5.3).

### Pourcentage d'erreur

On observe également un effet de compatibilité sur le pourcentage d'erreur ( $F_{(1,15)} = 30.3, p = 0$ ). Le pourcentage d'erreur est plus élevé lorsque l'essai est incompatible que lorsqu'il est compatible (6.1% versus 3.2%, voir figure 5.3).

### 5.2.4 Résultats Électrophysiologiques

Deux sujets parmi les seize présentaient des activités au niveau des électrodes d'intérêt ne nous permettant pas de faire correctement nos mesures. En effet les activités au niveau des aires motrices (électrodes C3, C4 et FCz) ne laissaient pas apparaître les potentiels d'intérêt. Nous avons choisi de les écarter pour l'ensemble des analyses statistiques sur les données électrophysiologiques. Elles ont donc été réalisées sur 14 sujets.

#### La négativité précédant le début de l'EMG (N-40)

Sur l'électrode FCz, nous avons analysé la négativité précédant le début de la réponse EMG. Le grand moyennage et la topographie de cette onde sont présentés sur la figure 5.4. Notons que cette activité est très localisée au niveau de l'électrode FCz. À partir du grand moyennage, nous observons un potentiel négatif précédant le début de l'EMG dans la condition incompatible (trait épais). En revanche, dans la condition compatible la négativité est très faible voire inexistante (trait fin). Les statistiques ont été réalisées à partir des tracés moyennés individuels. Nous avons d'abord analysé l'amplitude de l'onde, puis les pentes des courbes dans des fenêtres temporelles choisies sur la base du grand moyennage.

Pour mesurer l'amplitude de l'onde, nous avons repéré, dans l'intervalle de temps -150 0 ms, le pic le plus positif<sup>2</sup> puis le pic le plus négatif. Nous avons ensuite mesuré la différence d'amplitude entre les deux pics. L'analyse statistique a confirmé l'observation faite sur le grand moyennage. La négativité est plus grande dans la condition incompatible que dans la condition compatible ( $-0.19 \mu V/cm^2$  versus  $-0.15 \mu V/cm^2$ ;  $F_{(1,13)} = 5.8, p < 0.05$ ).

Les pentes de l'activité enregistrée en FCz ont été mesurées dans trois fenêtres temporelles de pente allant de -100 à -75 ms, de -75 à -35 ms et de -35 à 0 ms par rapport au début de l'activité EMG. Entre -100 et -75 ms, les pentes ne sont pas différentes de zéro quelle que soit la compatibilité au stimulus ( $t_{(13)} < 1$ ). Entre -75 et -35 ms, la pente dans la condition incompatible est négative et significativement différente de zéro ( $t_{(13)} = 6.6, p < 0.05$ ), alors que celle dans la condition

2. Rappelons que par convention, la négativité est représentée vers le haut

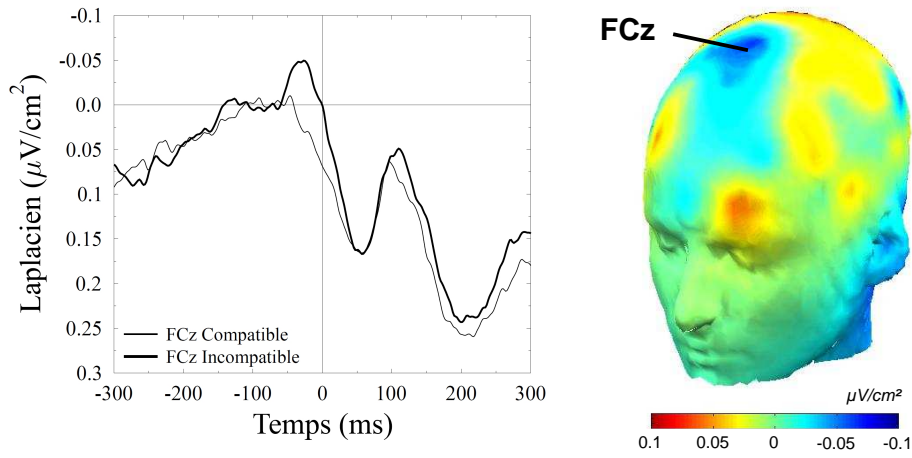


FIGURE 5.4 – **Expérience 1 : Dynamique temporelle et topographie de la N-40.** Les tracés fin et épais correspondent respectivement au grand moyennage de l'activité enregistrée dans les essais compatibles et incompatibles au niveau de l'électrode FCz. La carte située à droite représente la distribution du Laplacien de Surface 45 ms avant le début de l'EMG. On observe clairement la distribution fronto-centrale de cette onde, avec un maximum en FCz (électrode localisée sous le foyer bleu).

compatible n'est pas différence de zéro ( $t_{(13)} < 1$ ). La pente dans la condition incompatible est plus négative que celle dans la condition compatible (-0.15 versus -0.005,  $F_{(1,13)} = 17.9, p < 0.001$ ). Dans la fenêtre temporelle allant de -35 à 0 ms, la pente dans la condition incompatible est positive et différente de zéro ( $t_{(13)} = 9.4, p < 0.01$ ). Celle dans la condition compatible est positive mais l'effet statistique mesuré est tendanciel ( $t_{(13)} = 3.8, p = 0.075$ ). Les pentes dans les conditions compatible et incompatible ne sont pas différentes l'une de l'autre (0.12 versus 0.16,  $F_{(1,13)} < 1$ ).

Les résultats statistiques ont confirmé la présence d'une onde négative précédant le début de la réponse musculaire dans le condition incompatible. Elle semble se développer aux alentours de -75 ms, son maximum est enregistré aux alentours de -35 ms. Cette activité n'est pas observée dans la condition compatible.

#### Les activités au dessus des cortex moteurs :

Nous retrouvons bien le patron d'activation-inhibition décrit précédemment. Le grand moyennage des tracés au dessus des cortex moteur controlatéral et ipsilatéral au mouvement (respectivement en bleu et en rouge) ainsi que la topographie de ces activités sont présentés sur la figure 5.5. Au dessus du cortex moteur primaire controlatéral au mouvement (partie gauche de la tête sur la topographie), nous observons une négativité qui se développe environ 50 ms avant le début de l'EMG. Cette négativité, qui reflète probablement l'activation de la réponse, apparaît plus ample

dans la condition incompatible (trait bleu épais) que compatible (trait bleu fin). Au dessus du cortex moteur primaire ipsilatéral au mouvement effectué (partie droite de la tête sur la topographie), nous observons une positivité qui démarre aux alentours de 100 ms avant le début de l'EMG et qui ne semble pas être sensible à la compatibilité du stimulus (traits rouges). Comme pour la N-40, nous avons analysé l'amplitude des activités et ainsi que les pentes dans les fenêtres temporelles choisies sur la base du grand moyennage.

**Analyse de l'activation :** Pour mesurer l'amplitude de l'activation, nous avons repéré le pic minimum dans l'intervalle de temps entre -150 et 0 ms ainsi que le pic maximum dans l'intervalle de temps entre -50 à 100 ms, puis nous avons calculé la différence entre les deux amplitudes. L'observation des tracés moyennés laisse penser que l'activation dans les deux conditions démarre au même moment et que les deux courbes se distinguent quelques millisecondes après. Afin de tester au mieux cette hypothèse, nous avons choisi de mesurer les pentes dans des fenêtres d'analyse définies sur la base des tracés issus du grand moyennage.

L'analyse statistique confirme que l'amplitude de l'activation est plus grande dans la condition incompatible que dans la condition compatible ( $-0.31 \mu V/cm^2$  versus  $-0.26 \mu V/cm^2$ ,  $F_{(1,13)} = 5.1, p < 0.05$ ). Nous avons également testé l'effet de compatibilité sur la latence du pic. L'analyse statistique n'a pas révélé de différence significative de latence entre les deux pics (19 ms versus 18 ms,  $F_{(1,13)} < 1$ )

Nous avons analysé les pentes de l'activation dans les fenêtres temporelles allant de -75 à -45 ms, de -45 à -20 ms, de -20 à 5 ms et de 5 à 75 ms par rapport au début de l'activité EMG.

Entre -75 et -45 ms, les pentes ne sont différentes de zéro ni dans la condition compatible ( $t_{(13)} = 0.4, p = 0.55$ ), ni dans la condition incompatible ( $t_{(13)} = 2.7, p = 0.13$ ). Il n'y a pas de différence significative entre la valeur des pentes dans les deux conditions ( $F_{(1,13)} = 2.1, p = 0.17$ ). Entre -45 et -20 ms, les pentes dans les conditions compatible et incompatible sont toutes deux négatives et significativement différentes de zéro ( $t_{(13)} = 6.5, p < 0.05$  et  $t_{(13)} = 11.1, p < 0.01$ ). Pour autant, il n'y a pas de différence significative entre les valeurs des pentes entre les conditions compatible et incompatible ( $-0.25$  versus  $-0.31$ ,  $F_{(1,13)} < 1$ ). Entre -20 et 5 ms, les pentes dans les conditions compatible et incompatible sont toutes deux négatives et significativement différentes de zéro ( $t_{(13)} = 5.2, p < 0.05$  et  $t_{(13)} = 8.3, p < 0.05$ ). Par ailleurs, la pente est plus raide dans la condition incompatible que dans la condition compatible ( $-0.41$  versus  $-0.24$ ,  $F_{(1,13)} = 5.4, p < 0.05$ ). Enfin entre 5 et 75 ms, les pentes dans les conditions compatible et incompatible sont toutes deux positives et significativement différentes de zéro ( $t_{(13)} = 11, p < 0.01$  et  $t_{(13)} = 11, p < 0.01$ ).

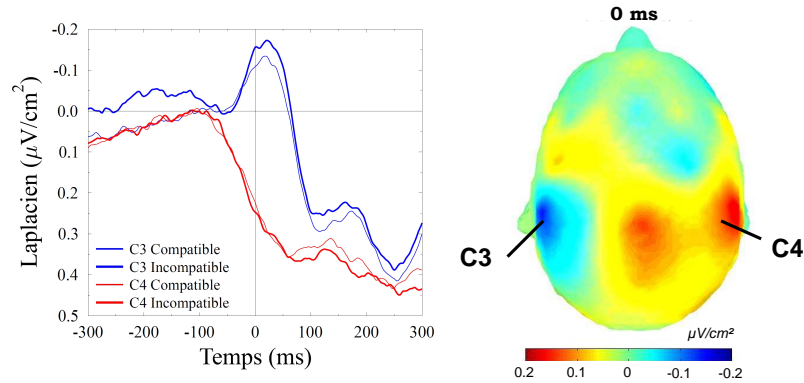


FIGURE 5.5 – **Expérience 1 : Sensibilité à la compatibilité du patron d'activation-inhibition obtenu au dessus des cortex moteurs M1 ipsilatéral et controlatéral au mouvement.** Les tracés bleu et rouge correspondent au grand moyennage de l'activité enregistrée respectivement au dessus de M1 controlatéral et ipsilatéral à la réponse. Les traits fins et épais correspondent respectivement à un stimulus compatible et incompatible. Le zéro des temps correspond au début de l'EMG. Cette carte représente la distribution du Laplacien de Surface au moment du début de l'activité EMG. On observe clairement une zone bleu caractéristique de l'activation au dessus du cortex moteur de l'hémisphère controlatéral au mouvement (électrode C3) et une zone rouge caractéristique de l'inhibition au dessus du cortex moteur situé sur l'hémisphère ipsilatéral au mouvement (électrode C4).

La raideur de la pente ne varie pas entre les conditions compatible et incompatible (0.33 versus 0.32,  $F_{(1,13)} < 1$ ).

**Analyse de l'inhibition :** Nous avons analysé les pentes de l'activité enregistrée au dessus du cortex moteur ipsilatéral au mouvement dans les fenêtres temporelles de pente allant de -100 à -75 ms, -75 à -25 ms et de -25 à 50 ms par rapport au début de l'activité EMG. Entre -100 et -75 ms, les pentes dans les deux conditions ne sont pas différentes de zéro ( $t_{(13)} < 1$ ). Aucune différence significative n'a été révélée par la comparaison des pentes dans les deux conditions ( $F_{(1,13)} < 1$ ).

Dans la fenêtre temporelle allant de -75 à -25 ms, les pentes des essais compatibles et incompatibles sont positives et significativement différentes de zéro ( $t_{(13)} = 12.7, p < 0.01$  et  $t_{(13)} = 18.5, p < 0.001$ ). Là non plus, il n'y a pas de différence significative entre les valeurs des pentes dans les conditions compatible et incompatible (0.22 versus 0.28,  $F_{(1,13)} = 1.6, p = 0.23$ ). Les analyses dans la fenêtre allant de -25 à 50 ms vont dans le même sens. Les pentes sont positives et différentes de zéro dans les deux conditions ( $t_{(13)} = 47.3, p < 0.001$  et  $t_{(13)} = 26.3, p < 0.001$ ). Encore une fois, la raideur de la pente n'est pas sensible à la compatibilité du stimulus ( $F_{(1,13)} < 1$ ).



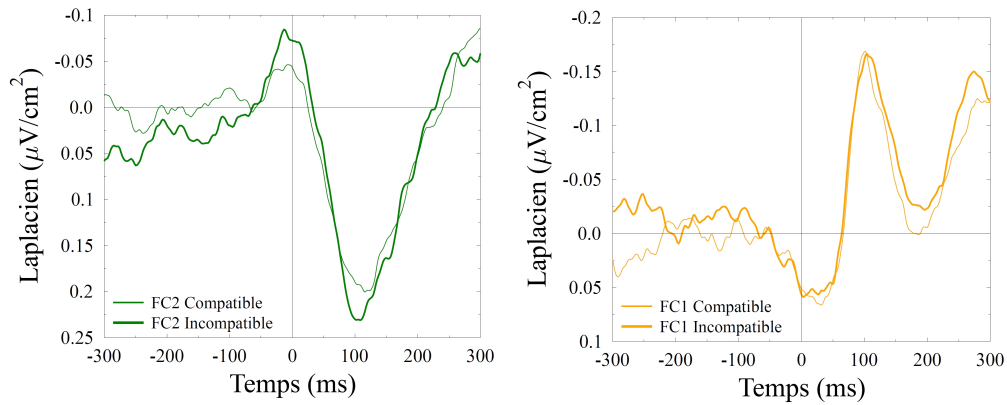


FIGURE 5.6 – **Expérience 1 : Dynamique temporelle de l'activité EEG au dessus des cortex prémoteurs calée sur le début de l'EMG en fonction de la compatibilité du stimulus.** Les traits verts et oranges correspondent respectivement aux grands moyennages de l'activité enregistrée au dessus des cortex pré-moteur ipsilatéral et controlatéral au mouvement. Les traits fins et épais correspondent respectivement aux essais dans lesquels le stimulus était compatible et aux essais dans lesquels le stimulus était incompatible. Le zéro des temps indique le début de l'activité EMG.

### Réponse électromyographique :

Nous avons testé la sensibilité de l'effet de la compatibilité sur la bouffée EMG. L'activité EMG semble être la même quelle que soit la nature du stimulus. Après avoir redressé le signal, nous avons mesuré l'amplitude de la bouffée EMG en relevant la valeur (en unité arbitraire) du pic maximum de la bouffée EMG dans les deux conditions. L'analyse a révélé que l'amplitude de la bouffée EMG n'était pas sensible à la compatibilité du stimulus ( $F_{(1,13)} < 1$ ).

### Activité au dessus des cortex prémoteurs :

Au niveau du cortex pré-moteur ipsilatéral à la réponse, nous remarquons une activité négative susceptible d'avoir un rôle dans l'organisation de la commande motrice. Elle se développe environ 100 avant le début de l'EMG et semble sensible à la compatibilité (tracé vert sur la figure 5.6). Aucune activité de ce type n'est observée au dessus du cortex controlatéral à la même latence (tracé orange sur la figure 5.6).

Nous avons réalisé des analyses statistiques permettant de tester la sensibilité de cette négativité à la compatibilité en mesurant son amplitude dans les deux conditions ainsi que les pentes dans les fenêtres temporelles choisies sur la base des tracés du grand moyennage.

L'analyse de l'amplitude a révélé que cette négativité est plus ample dans la condition incompatible que dans la condition compatible ( $-0.25 \mu V/cm^2$  versus  $-0.19 \mu V/cm^2$ ,  $F_{(1,13)} = 5.2, p < 0.05$ ).

Nous avons analysé les pentes dans des fenêtres temporelles allant de -75 à -60 ms, de -60 à -30 ms, de -30 à -15 ms et de -15 à 50 ms.

Entre -75 et -60 ms, l'analyse des pentes ne révèle que les pentes dans les conditions compatible et incompatible ne sont pas différentes de zéro ( $F_{(1,13)} = 1.4, p = 0.26$  et  $F_{(1,13)} = 1.7, p = 0.22$ ). En revanche, une différence marginale est observée entre la raideur de la pente dans la condition compatible et la condition incompatible (0.12 versus -0.12,  $p < 1.10$ ). Entre -60 à -30 ms, les pentes dans les conditions compatible et incompatible sont négatives et différentes de zéro ( $t_{(13)} = 7.9, p < 0.05$  et  $t_{(13)} = 5.2, p < 0.05$ ). Par ailleurs, aucune différence de pente n'est observée entre les deux conditions (-0.15 versus -0.16,  $F_{(1,13)} < 1$ ). Dans la fenêtre temporelle allant de -15 à 50 ms, les pentes dans les deux conditions sont positives et différentes de zéro ( $t_{(13)} = 17.3, p < 0.01$  et  $t_{(13)} = 9.2, p < 0.01$ ). Aucune différence de pente entre les deux conditions n'est révélée par l'analyse ( $F_{(1,13)} < 1$ ).

Cette négativité prémotrice est spécifique de l'hémisphère ipsilatérale à la réponse. En effet, au niveau des électrodes équivalent contralatérale, aucune négativité n'est observable, et l'on voit au contraire une légère positivité.

### **Hierarchie entre la N-40, la négativité prémotrice ipsilatérale et l'activation**

Cette activité ipsilatérale semble se développer entre la négativité enregistrée en FCz et l'activation enregistrée en C3. Nous avons donc testé les latences des pics maximum des trois ondes à partir des tracés moyennés de chaque sujet dans lesquels les essais compatibles et incompatibles étaient mélangés.

La latence du pic maximum de la négativité ipsilatérale prémotrice se situe entre le maximum du pic de l'onde de la N-40 et de celui de l'activation. Les latences moyennes respectives des trois ondes négatives sont respectivement -45.2 ms, -12.4 ms et 12.9 ms pour la N-40, la négativité prémotrice ipsilatérale et l'activation. L'analyse statistique montre que la latence du pic de la N-40 précède la latence du pic de la négativité prémotrice ( $F_{(1,13)} = 14.8, p < 0.01$ ); et la latence de la négativité prémotrice ipsilatérale précède la latence du pic de l'activation enregistrée sur M1 contralatéral au mouvement ( $F_{(1,13)} = 4.6, p = 0.05$ ).

#### **5.2.5 Discussion de l'expérience 1 :**

Afin de préciser les mécanismes de sélection de la réponse, nous avons étudié les activités prefrontales et motrices qui se mettent en place juste avant l'émission de la réponse dans une tâche

de TR de choix impliquant un conflit cognitif. Nous retrouvons :

- Une activité fronto-centrale (FCz, au-dessus de l'AMS, décrite par Vidal *et al.*, 2003b). Cette activité est plus ample pour les essais incompatibles
- le patron d'activation-inhibition des réponses déjà rapporté (Taniguchi *et al.*, 2001; Tandonnet *et al.*, 2003; Vidal *et al.*, 2003b). L'activation est plus large pour les essais incompatible. L'inhibition n'est pas affectée.
- L'onde fronto-centrale précède les activités des cortex moteurs (Vidal *et al.*, 2003b).

De plus, nous rapportons une onde négative, ipsilatérale à la réponse, plus ample en incompatible, et dont le décours temporel est entre AMS et M1. Un résumé des effets obtenus est illustré dans la figure 5.7. Afin de préciser et de mieux établir ces faits, nous avons cherché à les répliquer dans une autre tâche : tâche d'Eriksen.

## 5.3 Expérience 2

### 5.3.1 Introduction

Afin de valider et de répliquer les données obtenues dans la tâche de Simon, nous avons effectué les mêmes analyses sur un jeu de données obtenu dans une autre tâche de conflit : la tâche d'Eriksen (cf. chapitres expérimentaux précédents). L'objectif était d'étudier si les différences observées dans la tâche de Simon étaient spécifiques à cette tâche, ou si au contraire ils reflétaient des mécanismes généraux liés à la sélection de la réponse et à l'organisation de la commande motrice.

Pour ce faire, nous avons re-analysé les données issues de l'expérience sur la tâche d'Eriksen afin de regarder les mêmes composants électrophysiologiques que ceux analysés dans l'expérience que nous venons de présenter.

### 5.3.2 Matériel - Méthode

Ces données ayant déjà été décrite précédemment, nous rappelons brièvement la méthode utilisée.

Dix sujets ont participé à l'expérience. Ils devaient réaliser une tâche d'Eriksen (Eriksen et Eriksen, 1974). Les stimuli étaient composés de trois lettres contiguës. Les sujets avaient comme consigne de répondre le plus vite possible en fonction de la nature de la lettre centrale en appuyant sur le bouton placé dans la main gauche ou sur celui placé dans la main droite. Les première et troisième lettres sont appelées distracteurs. Dans un cas, les trois lettres étaient identiques,

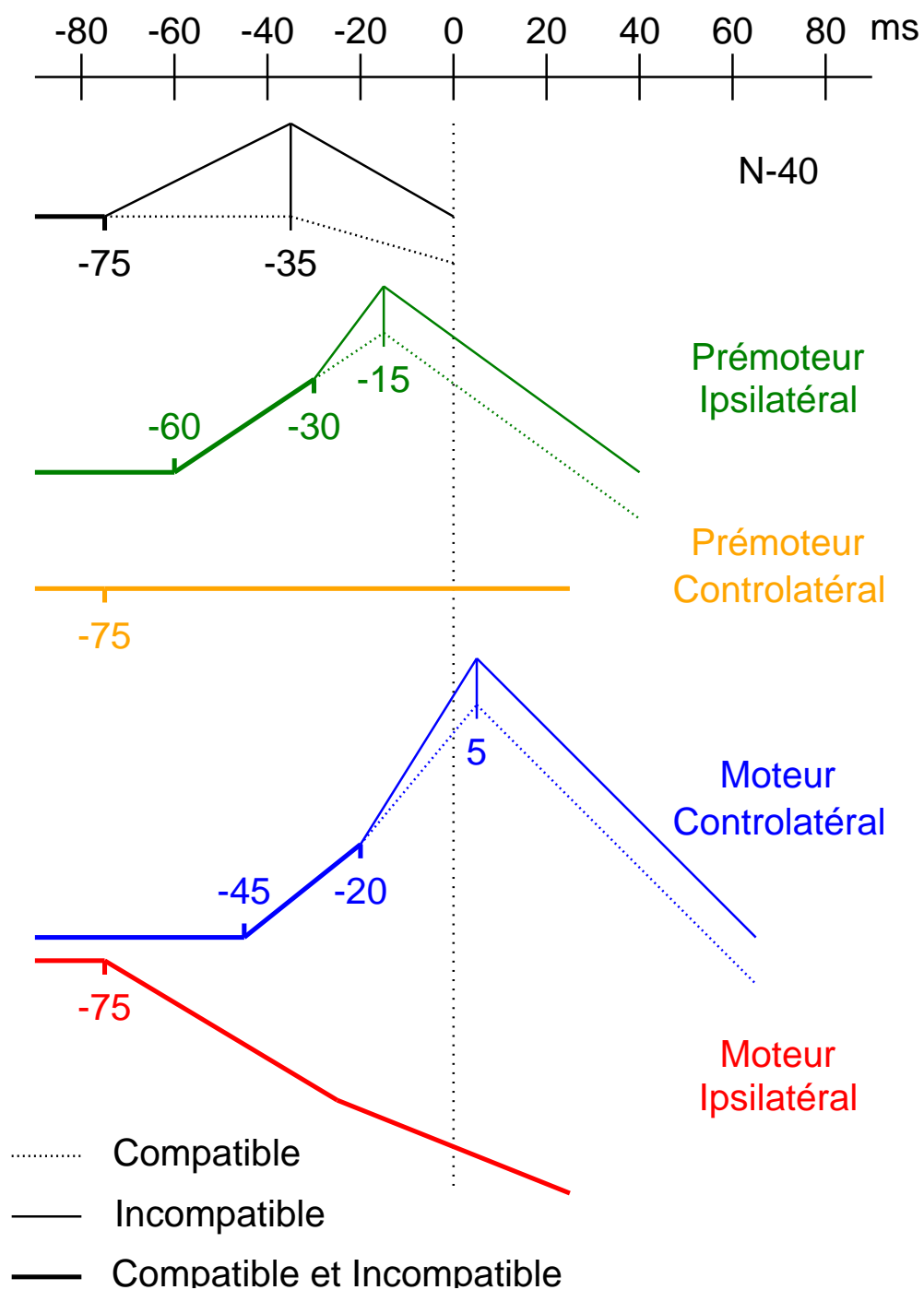


FIGURE 5.7 – Expérience 1 : Dynamique temporelle des activations dans les régions motrices aux alentours de la réponse.

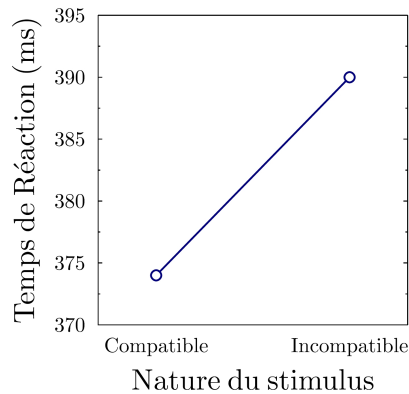


FIGURE 5.8 – **Expérience 2 : Résultats Comportementaux.** Temps de réaction moyens chez les sujets en fonction de la compatibilité du stimulus.

le stimulus était “compatible” (stimuli SSS et HHH). Dans l’autre cas, les lettres distractrices pouvaient être différentes de la lettre centrale, le stimulus était “incompatible” (stimuli HSH et SHS). Dans cette expérience, les sujets réalisaient 2560 essais, répartis en 20 blocs de 128 essais. Les sujets disposaient de 1000 ms pour répondre. Le délai entre la réponse et le stimulus de l’essai suivant était de 1000 ms.

Le recueil des données a été réalisé dans des conditions comparables et avec le même matériel que dans l’expérience précédente. Les étapes de traitement des données EEG et EMG sont également identiques à celles décrites dans l’expérience précédente.

Pour chacun des 10 sujets, nous avons sélectionné tous les essais purs-corrects et analysé les enregistrements EEG au niveau des aires motrices primaires (C3 et C4), au niveau des aires prémotrices (FCz, FC1 et FC2), ainsi que la bouffée EMG en fonction de la compatibilité du stimulus. Toutes nos analyses EEG ont été réalisées après calcul du Laplacien de surface sur les données EEG monopolaires.

### 5.3.3 Résultats comportementaux

Les pourcentages d’essais purs-corrects, corrects avec ébauche d’erreur et erreurs sont respectivement de 76.8%, de 18% et de 5.2%.

On observe un effet de compatibilité sur le TR des essais pur-corrects. Les TR dans les essais compatibles sont significativement plus courts que les TR dans les essais incompatibles (374 ms versus 390 ms,  $F_{(1,9)} = 93.5, p < 0.001$ ).

### 5.3.4 Résultats Électrophysiologiques

Comme dans l'expérience précédente, nous avons analysé la sensibilité à la compatibilité des activités enregistrées dans les aires motrices calées sur le début de la réponse EMG. Plus précisément, nous nous sommes intéressés à la N-40, au patron d'activation-inhibition, aux activités prémotrices, ainsi qu'à la forme de la bouffée électromyographique.

#### La négativité précédant le début de l'EMG (N-40)

La N-40 apparaît très clairement sur le grand-moyennage de l'électrode FCz, et semble être plus grande dans la condition incompatible. Contrairement aux résultats de l'expérience précédente, la N-40 semble également être présente dans les essais compatibles (voir figure 5.9). Pour valider statistiquement ces observations, nous avons comparé l'amplitude de l'onde dans les deux conditions. Nous avons ensuite mesuré les pentes dans des fenêtres temporelles choisies sur la base du grand moyennage.

Comme dans l'expérience précédente, nous avons mesuré l'amplitude pic à pic de cette onde dans les essais pur-corrects. Dans la fenêtre de temps -150 0 ms, nous avons identifié le pic minimum puis le pic maximum de l'onde puis nous avons mesuré la différence d'amplitude entre les deux pics.

L'analyse statistique a confirmé que l'amplitude de l'onde est plus grande en situation incompatible que compatible ( $-0.26 \mu V/cm^2$  versus  $-0.19 \mu V/cm^2$ ,  $F_{(1,9)} = 37.4, p < 0.001$ ). La différence de latence du pic maximum dans les conditions compatible et incompatible n'est pas significative ( $-34.5$  ms versus  $-22.2$  ms,  $F_{(1,9)} = 3.1, p = 0.11$ ).

Nous avons analysé les pentes de ces activités dans les fenêtres temporelles de pente allant de -120 à -90 ms, de -90 à -60 ms, de -60 à -30 ms et de -30 à 70 ms par rapport au début de l'activité EMG. Entre -120 et -90 ms, les pentes ne sont pas différentes de zéro quelle que soit la compatibilité  $S-R(t_{(9)} < 1)$ . Entre -90 et -60 ms, la pente dans la condition incompatible est négative et différente de zéro ( $t_{(9)} = 7.7, p < 0.05$ ), alors que celle dans la condition compatible n'est pas différente de zéro ( $t_{(9)} = 1.2, p = 0.29$ ). L'analyse statistique révèle une différence marginale entre les pentes dans les deux conditions ( $0.07$  versus  $-0.18$ ,  $F_{(1,9)} = 4.6, p = 0.06$ ). Entre -60 et -30 ms, les pentes sont négatives et significativement différentes de zéro ( $t_{(9)} = 12.2, p < 0.01$  et  $t_{(9)} = 7.3, p < 0.05$ ). La raideur de la pente n'est pas influencée par la compatibilité ( $F_{(1,9)} < 1$ ). Entre -30 et 70 ms, la pente est positive et différente de zéro dans la condition incompatible ( $t_{(9)} = 5.9, p < 0.05$ ), tandis que celle dans la condition compatible n'est que marginalement

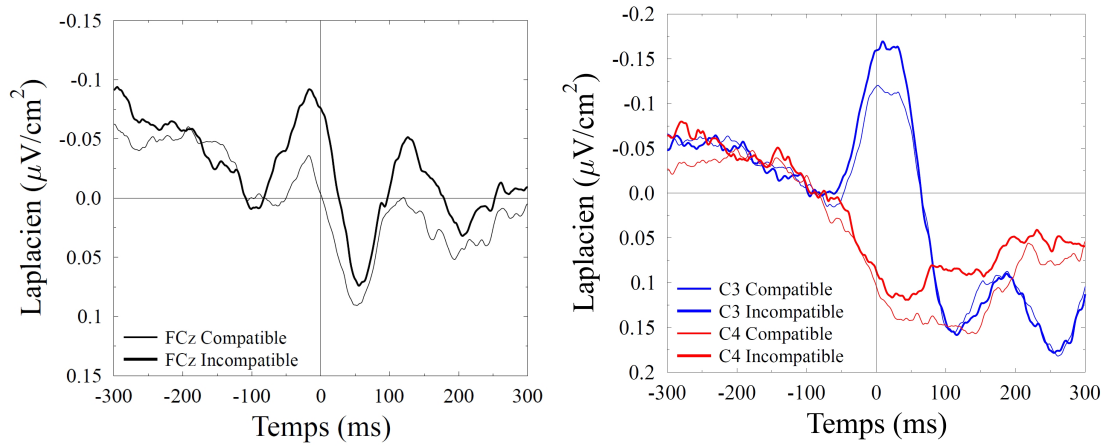


FIGURE 5.9 – **Expérience 2 : Dynamique temporelle de la N-40 et du patron d’activation-inhibition.** Les tracés fin et épais correspondent respectivement au grand moyennage de l’activité enregistrée dans les essais compatibles et incompatibles au niveau de l’électrode FCz (en noir), du cortex M1 controlatéral au mouvement avec l’électrode C3 (en bleu) et du cortex M1 ipsilatéral au mouvement avec l’électrode C4 (en rouge). Le zéro des temps correspond au début de l’activité EMG.

différente de zéro ( $t_{(9)} = 3.6, p = 0.08$ ). L’effet de la compatibilité sur les pentes dans cette fenêtre temporelle montre un effet marginal ( $F_{(1,9)} = 4.2, p = 0.07$ ). Il semblerait que la pente soit davantage positive dans la condition incompatible que dans la condition compatible (0.16 versus 0.11).

#### Activités au dessus des cortex moteurs :

Nous retrouvons bien le patron d’activation-inhibition décrit précédemment, et comme dans la version accessoire de la tâche de Simon, l’activation semble être sensible à la compatibilité (voir figure 5.9).

**Analyse de l’activation :** L’observation des tracés moyennés laisse penser que l’activation dans les deux conditions démarre au même moment et que les deux courbes se distinguent quelques millisecondes après. Afin de tester au mieux cette hypothèse, nous avons mesuré l’amplitude de l’activation dans les deux conditions, puis choisi de tester les pentes dans des fenêtres d’analyse définies sur la base des tracés issus du grand moyennage.

L’amplitude de l’activation est plus grande dans la condition incompatible que compatible ( $-0.34 \mu V/cm^2$  versus  $-0.27 \mu V/cm^2$ ,  $F_{(1,9)} = 5.1, p < 0.05$ ). De plus, la compatibilité du stimulus n’influe pas sur la latence du pic maximum de l’activation ( $F_{(1,9)} < 1$ ). La latence moyenne du pic

maximum de l'activation, dans les conditions compatible et incompatible, se situe respectivement à 18 et 19 ms après le début de l'EMG.

Nous avons analysé les pentes de l'activation dans les fenêtres temporelles allant de -70 à -50 ms, de -50 à -0 ms et de 0 à 75 ms par rapport au début de l'activité EMG.

Entre -70 à -50 ms, les pentes ne sont pas différentes de zéro dans les conditions compatible et incompatible ( $t_{(9)} < 1$ ). Entre -50 à 0 ms, les courbes sont négatives et différentes de zéro ( $t_{(9)} = 27.6, p < 0.001$  pour la condition compatible et  $t_{(9)} = 11.8, p < 0.01$  pour la condition incompatible). Mais l'analyse statistique ne montre pas de différence de pente entre les conditions compatible et incompatible (-0.28 versus -0.34,  $F_{(1,9)} < 1$ ). Entre 0 et 75 ms après le début de l'EMG, les pentes sont positives et différentes de zéro dans les deux conditions ( $t_{(9)} = 6.5, p < 0.05$  et  $t_{(9)} = 9.4, p < 0.05$ ). L'effet de compatibilité sur la pente est marginal ( $F_{(1,9)} = 3.4, p = 0.10$ ). Les pentes apparaissent plus positives dans la condition incompatible que dans la condition compatible (0.30 versus 0.21). À l'issue de ces analyses, nous n'avons pas réussi à mettre en évidence des différences significatives entre les pentes dans les conditions compatible et incompatible.

**Analyse de l'inhibition :** Dans la fenêtre temporelle allant de -90 à 0 ms, les pentes de l'inhibition dans les conditions compatible et incompatible sont positives et différentes de zéro ( $t_{(9)} = 7.7, p < 0.05$  et  $t_{(9)} = 8.0, p < 0.05$ ). Aucune différence significative entre les deux courbes n'est observée ( $F_{(1,9)} < 1$ ).

#### **Amplitude de la réponse électromyographique :**

Nous avons testé la sensibilité de la bouffée EMG à la compatibilité du stimulus. L'activité EMG semble être la même quelle que soit la nature du stimulus. Après avoir redressé et intégré le signal, nous avons mesuré l'amplitude de la bouffée EMG en relevant la valeur (en unité arbitraire) du pic maximum de la bouffée EMG dans les deux conditions. Comme dans l'expérience de précédente, l'analyse a révélé que l'amplitude de la bouffée EMG n'était pas sensible à la compatibilité du stimulus ( $F_{(1,9)} < 1$ ).

#### **Activité au dessus des cortex prémoteurs :**

Après observation des activités présentes au dessus des cortex prémoteurs, il ne nous semble pas que la négativité prémotrice ipsilatérale précèdent le début de l'EMG soit sensible à la compatibilité. Aucune activité particulière ne semble être présente au dessus du cortex pré moteur contralatéral au mouvement avant le début de l'EMG.



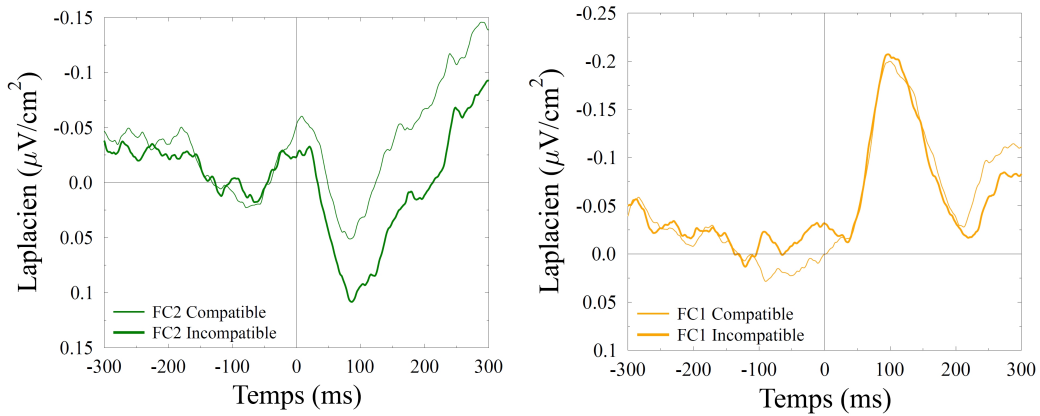


FIGURE 5.10 – **Expérience 2 : Dynamique temporelle de l’activité EEG au dessus des cortex prémoteurs calée sur le début de l’EMG en fonction de la compatibilité du stimulus.** Les traits verts et oranges correspondent respectivement aux grands moyennages de l’activité enregistrée au dessus des cortex pré-moteur ipsilatéral et contralatéral au mouvement. Les traits fins et épais correspondent respectivement aux essais dans lesquels le stimulus était compatible et aux essais dans lesquels le stimulus était incompatible. Le zéro des temps indique le début de l’activité EMG.

Nous avons analysé les pentes dans les fenêtres temporelles allant de -90 à -70 ms, de -70 à -10 ms et de -10 à 75 ms. Dans la première fenêtre temporelle, les pentes dans les conditions compatible et incompatible ne sont pas différentes de zéro ( $t_{(9)} < 1$ ). Dans la fenêtre temporelle allant de -70 à -10 ms, les pentes sont significativement négatives ( $t_{(9)} = 15.9, p < 0.001$  pour la condition compatible et  $t_{(9)} = 6.3, p < 0.05$  pour la condition incompatible). En revanche, il n’y a pas de différence significative entre les deux courbes ( $F_{(1,9)} < 1$ ). Dans la fenêtre temporelle allant de -10 à 75 ms, les pentes dans les deux conditions sont positives ( $t_{(9)} = 13.3, p < 0.001$  pour la condition compatible et  $t_{(9)} = 19.7, p < 0.001$  pour la condition incompatible). Là encore, il n’y a pas de différence significative entre ces deux conditions ( $F_{(1,9)} = 3.24, p = 0.1$ ). L’analyse de l’amplitude de cette onde ne révèle pas de différence significative d’amplitude entre les conditions compatible et incompatible ( $-0.18 \mu V/cm^2$  versus  $-0.19 \mu V/cm^2$ ,  $F_{(1,9)} < 1$ ). Comme pour l’expérience précédente, cette négativité pré-motrice est restreinte à l’hémisphère ipsilatéral à la réponse. En effet, aucune onde, ni aucune différence significative entre compatible et incompatible n’est observée sur les électrodes contralatérales ( $F_{(1,9)} \max = 2.6, p = .14$ ).

#### Hierarchie entre la N-40, la négativité pré-motrice ipsilatérale et l’activation

L’analyse des latences des pics maximum des activités négatives enregistrées en FCz, FC2 et C3 révèle une différence significative entre le pic maximum de la N-40 et celui de l’activité

prémotrice ipsilatérale (-22.3 ms versus 3.3 ms,  $F_{(1,9)} = 9.3, p < 0.05$ ), ainsi qu'entre celui de l'activité prémotrice ipsilatérale et le pic maximum de l'activation (3.3 ms versus 25.2 ms,  $F_{(1,9)} = 9.1, p < .05$ ).

### 5.3.5 Discussion de l'expérience 2 :

Nous retrouvons bien les effets obtenus dans l'expérience précédente, à savoir :

- La N-40 est sensible à la compatibilité. Son amplitude est plus importante dans la condition incompatible qu'en condition compatible.
- Sur le patron d'activation-inhibition, l'activation observée sur le cortex M1 contralatéral à la réponse est plus importante dans la condition incompatible. L'inhibition observée sur le cortex M1 ipsilatéral à la réponse n'est pas sensible à la compatibilité.
- Nous retrouvons la négativité intermédiaire au niveau prémoteur ipsilatéral. Mais, contrairement aux résultats de l'expérience précédente, cette activité n'est pas sensible à la négativité.

Ces différents points sont discutés dans la discussion générale.

## 5.4 Discussion des deux expériences

Le but de l'étude était de chercher à préciser les mécanismes physiologiques de la prise de décision. Nous avons distingué deux grands types de modèles computationnels visant à rendre compte des mécanismes de sélection de la réponse. Une première catégorie de modèles ne prévoit pas d'étape S-R, la sélection de la réponse étant effectuée directement par les unités chargées d'exécuter la réponse. Les modèles de la deuxième catégorie prévoient une étape d'association S-R indépendante de l'exécution des réponses.

Nous avons cherché à tester ces deux catégories de modèles en étudiant des activités EEG candidates à ces étapes : le patron d'activation-inhibition enregistré au dessus du cortex M1 et la N-40 enregistrée en regard de l'AMS. Les activités prémotrices ont également été étudiées. Si, comme le prévoit les modèles sans étape S-R, la sélection de la réponse est effectuée au sein des régions chargées de l'exécution de la réponse, alors nous devrions observer un patron d'activation-inhibition qui démarre plus tôt avant la réponse dans la condition incompatible que dans la condition compatible (Yeung *et al.*, 2004). En revanche, s'il existe une étape de sélection de la réponse indépendante de l'exécution de la réponse, la forme du patron d'activation-inhibition devrait être similaire dans les deux conditions. Dans ce cas de figure, une différence entre les deux conditions pourrait être observée au niveau de la N-40.

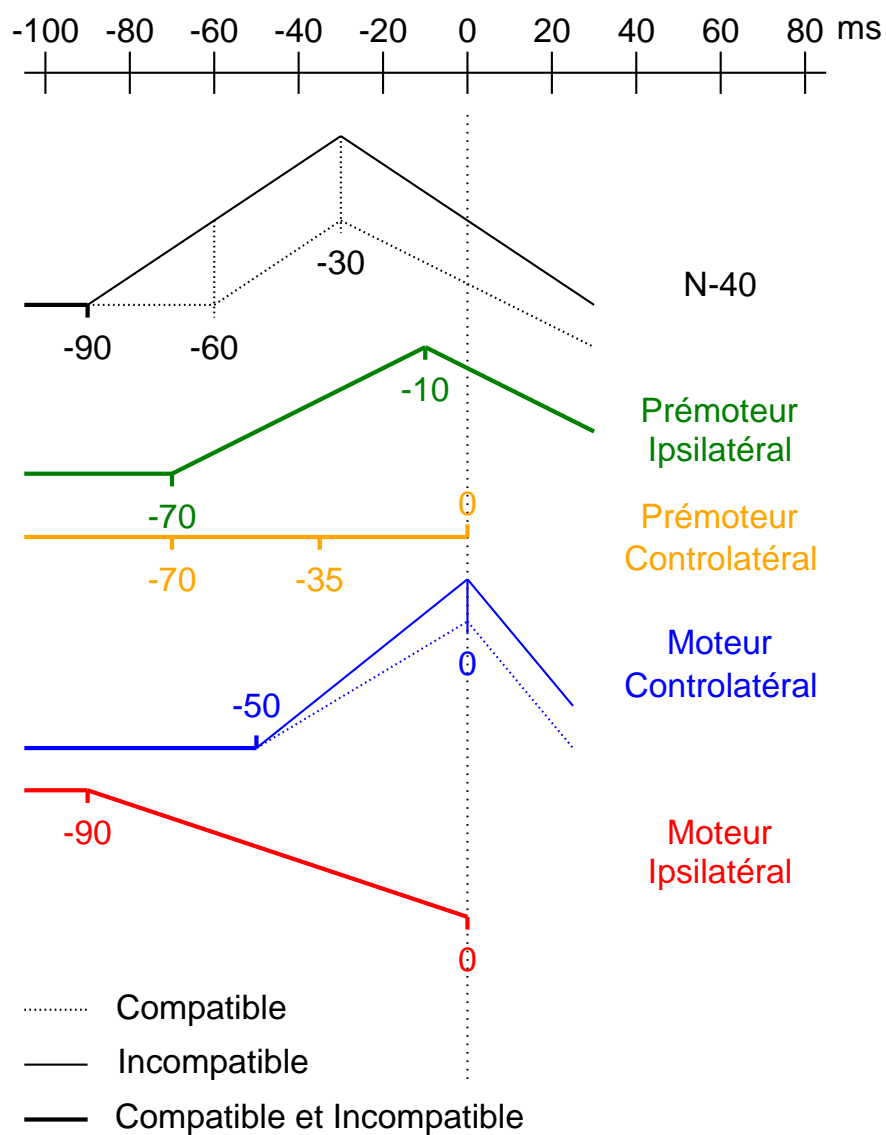


FIGURE 5.11 – Expérience 2 : Dynamique temporelle des activations dans les régions motrices aux alentours de la réponse.

Nous avons testé ces hypothèses dans deux jeux de données expérimentales recueillies chez des participants effectuant soit une adaptation de la tâche de Simon, soit une tâche d'Eriksen. Ces deux tâches diffèrent par la nature du conflit qu'elles impliquent (Kornblum *et al.*, 1990), ainsi que par les modalités perceptives qu'elles recrutent : dans la tâche de Simon, l'attribut pertinent était présenté visuellement, alors que l'attribut non-pertinent était auditif. Par contre, dans la tâche d'Eriksen, les attributs pertinents et non-pertinents étaient présentés dans la même modalité. La comparaison des résultats dans les deux expériences permet d'évaluer le caractère général ou spécifique (à une modalité de présentation ou encore à une tâche particulière) des effets obtenus.

Les résultats ont montré que les composantes observées étaient similaires dans les deux études. Nous retrouvons bien le patron d'activation-inhibition et la N-40. Une activité négative précédant la réponse a également été observée au dessus des aires prémotrices sur le cortex ipsilatéral à la réponse. Dans les deux expériences, la chronométrie entre les activités est la même. Les trois négativités sont séquencées dans le temps. La N-40 qui atteint son pic d'activité en premier, est suivie de l'activité prémotrice ipsilatérale à la réponse puis de l'activation au dessus de M1 contralatéral à la réponse. L'inhibition au dessus de M1 ipsilatérale à la réponse semble commencer avant l'activation. On note également qu'elle semble démarrer au même moment que la N-40. L'effet de la compatibilité du stimulus est similaire dans les deux expériences sur la N-40 et sur le patron d'activation-inhibition. La N-40 et l'activation au dessus de M1 contralatéral à la réponse sont de plus grandes amplitudes dans la condition incompatible que dans la condition compatible. L'inhibition n'est pas affectée par la compatibilité (voir figures 5.7 et 5.11). En revanche, la négativité prémotrice ipsilatérale à la réponse présente une sensibilité différente à la compatibilité dans les deux tâches : elle est plus grande dans la condition incompatible que dans la condition compatible dans l'expérience de Simon, alors qu'elle ne semble pas être sensible à la compatibilité dans la tâche d'Eriksen.

#### 5.4.1 Rôle des cortex moteurs primaires

Le mode de sélection de la réponse implémenté par le modèle du conflit (Botvinick *et al.*, 2001) émane des modèles de prise de décision synthétisés par Usher et McClelland (2001). Dans ces modèles, les unités de réponse en compétition accumulent des "preuves" en faveur de la réponse qu'ils représentent. Chaque unité de réponse exerce une inhibition, dite "latérale", sur l'unité de réponse alternative. La force de l'inhibition est proportionnelle au niveau d'activité de l'unité. La compétition entre les deux réponses est remportée par celle dont l'activité dépasse en premier le

seuil de réponse. La présence de l'inhibition latérale prédit que les TR dans les essais incompatibles sont plus longs que dans les essais compatibles, car dans les essais incompatibles, l'activation de la réponse incorrecte par les distracteurs inhibe la réponse correcte et l'empêche de franchir le seuil de réponse aussi rapidement que dans les essais compatibles. De plus, Yeung *et al.* (2004) ont comparé la dynamique d'activation des unités de réponse dans les essais compatibles et incompatibles. L'activation de l'unité de réponse correcte démarrait plus tôt avant la réponse dans la condition incompatible que dans la condition compatible (voir figure 4 page 138 de Yeung *et al.*, 2004). Si les cortex moteurs sont bien le reflet de l'activation des unités de réponse, alors nos résultats vont à l'encontre du modèle. En effet, si nous obtenons bien un effet sur l'activation de M1 controlatéral à la réponse, celui-ci est différent de l'effet prédit par la simulation du modèle. Dans les deux expériences, l'activation démarre en même temps dans les deux conditions et son amplitude est augmentée pour la condition incompatible.

Le fait que l'augmentation de l'activation au dessus de M1 controlatéral n'implique pas d'augmentation de la réponse EMG est surprenant. Cette observation suggère l'existence d'une modulation du seuil de réponse implémentée en cours d'essai au niveau des cortex moteurs primaires. Les modèles neuromimétiques utilisent souvent des fonctions à seuil pour modéliser la décharge de neurones ou des populations de neurones (voir figure 2.2 page 44). La modification du seuil de réponse en fonction d'un contexte est le procédé classiquement implémenté dans des modèles neuromimétiques pour rendre compte de l'effet de l'échange vitesse/précision sur la performance (Smith et Ratcliff, 2004; Yeung *et al.*, 2004).

Nous suggérons que l'augmentation de l'activation au dessus de M1 controlatéral dans la condition incompatible couplée à l'absence d'augmentation de l'activité EMG pourrait être expliquée par la modification du seuil de réponse en cours d'essai. Une telle augmentation pourrait être un moyen d'éviter les erreurs. Cette vision implique que l'activité au dessus des cortex M1 n'est pas linéairement liée à une activité EMG, et que les neurones pyramidaux de M1 soient informés par une structure en amont de la nécessité d'élever le seuil d'envoi de la commande motrice. Dans ce cas, le système nerveux doit pouvoir détecter un risque d'erreur et entraîner l'élévation du seuil de déclenchement de la réponse au niveau des cortex M1. Cette hypothèse est cohérente avec les résultats d'une étude IRMf récente (Eichele *et al.*, 2008).

Nos résultats suggèrent que la sélection de la réponse ne se fait pas par une simple accumulation d'informations au niveau des cortex moteurs. Toutefois, le rôle de ces derniers ne semble pas être restreint à l'exécution de la réponse. Nous proposons qu'ils soient impliqués dans l'implémentation d'un mécanisme de prévention de l'erreur à travers la mise en place d'une élévation du seuil de

déclenchement des réponses lorsqu'un risque d'erreur est présent. Il reste à déterminer comment le cerveau arrive à détecter en cours d'essai la présence d'un risque d'erreur.

### 5.4.2 Le rôle de la N-40

Dans les deux tâches, l'amplitude de la N-40 est plus importante dans les essais incompatibles comparés aux essais compatibles. Ce résultat est cohérent avec ceux de Ullsperger et von Cramon (2001) qui montrent une augmentation de l'activité dans la partie rostrale de l'aire motrice supplémentaire (*pre-SMA*) dans les essais incompatibles par rapport aux essais compatibles. L'antériorité du pic maximum d'activité de la N-40 par rapport à celui de l'activation, couplée à la sensibilité de ces deux activités à la compatibilité suggère que la N-40 influence l'activation et qu'elle pourrait être impliquée dans la sélection de la réponse. Or sélectionner la bonne réponse implique l'utilisation des règles d'association S-R. On peut imaginer que l'effort nécessaire à la sélection de la bonne réponse est plus important en situation incompatible que compatible. Si ce raisonnement est valide, alors la N-40 est un bon candidat à l'implémentation de ces règles d'association S-R puisque son amplitude est plus importante en situation incompatible que compatible. Nous observons une différence entre les deux tâches au niveau de la N-40. Celle-ci n'est pas visible dans la condition compatible dans la tâche de Simon, alors qu'elle l'est clairement dans la tâche d'Eriksen. Dans la tâche de Simon, l'association S-R pour le distracteur (la position du son et le côté de réponse) est naturelle, alors que celle pour la cible est arbitraire (une couleur et un côté de réponse). L'effort nécessaire à la transformation S-R arbitraire ne concerne qu'un seul attribut du stimulus dans la tâche de Simon alors qu'il concerne les deux dans la tâche d'Eriksen (une lettre et un côté de réponse pour la cible et les distracteurs). Il est donc possible que dans les essais compatibles, la sélection de la bonne réponse nécessite moins d'effort dans la tâche de Simon que dans la tâche d'Eriksen. Cette notion d'effort peut également expliquer pourquoi la N-40 n'est pas observée dans les tâches où les associations S-R sont simples ou facilitées par une information sur la réponse à donner (Burle *et al.*, 2004b; Meynier *et al.*, 2009; Carbonnell *et al.*, 2004).

Toutefois, il reste le point délicat de la hiérarchie entre la N-40 et l'inhibition. Les résultats des deux expériences suggèrent que l'inhibition démarre au même moment que la N-40. Si cette observation est vraie, cela implique que le choix de la réponse est d'ors et déjà effectué et que la N-40 ne peut donc pas refléter l'association S-R, ni la sélection de la réponse. Il est possible que l'inhibition soit pilotée par une autre structure. Les ganglions de la base pourraient être un bon candidat à l'envoi de cette commande. Dans l'autre cas de figure, l'inhibition pourrait être

pilotée par la N-40. Si le mécanisme de prévention de l'erreur évoqué au paragraphe précédent est implémenté au niveau du cortex moteur controlatéral (augmentation du seuil de déclenchement des réponses en cas de risque d'erreur), il n'est pas utile d'inhiber davantage le cortex impliqué dans la réponse incorrecte.

Nous devons nous assurer que les différences temporelles observées sont bien réelles et non pas le fruit d'un artefact de moyennage. Cependant, il est difficile de savoir laquelle des deux activités démarre en premier après moyennage des essais. En effet, si les activités d'intérêt ne sont pas tout à fait synchronisées sur le début de l'EMG, il y a de fortes chances que le début de l'activité sur le grand moyennage ne corresponde pas au début moyen de l'activité d'intérêt à travers les essais, mais plutôt à l'activité d'intérêt des essais en marge de la distribution (Callaway *et al.*, 1984; Burle *et al.*, 2008b). Nous devons également prêter attention à un éventuel recouvrement temporel des activités EEG<sup>3</sup>. Malheureusement, ces problèmes ne pourront être résolus que lorsque l'analyse des données EEG essais par essais sera possible.

### 5.4.3 La négativité prémotrice ipsilatérale à la réponse

C'est la première fois qu'une activité précédant la réponse au niveau des cortex prémoteurs est rapportée dans les tâches de TR. Cette activité ipsilatérale à la réponse, atteint son pic maximum d'activité entre celui de la N-40 et celui de l'activation. De façon très intéressante, elle est sensible à la compatibilité dans la tâche de Simon alors qu'elle ne l'est pas dans la tâche d'Eriksen. L'interprétation de cette activité n'est pas aisée. Puisque nous observons des différences entre les deux tâches, il est probable que son rôle soit lié à ce qui diffère entre elles : c'est-à-dire le statut de l'attribut non-pertinent. Le distracteur active une réponse de façon arbitraire dans la tâche d'Eriksen et de façon plus naturelle dans la tâche de Simon. Bien que nous n'ayons pas d'explication précise sur la nature de cette activité, nous supposons qu'elle soit liée au traitement et/ou à la gestion de l'attribut non-pertinent.

## 5.5 Conclusion

L'analyse des activités EEG dans les régions motrices dans les expériences de Simon et d'Eriksen montre de nombreuses similitudes. Au vu de la sensibilité du patron d'activation-inhibition à la compatibilité du stimulus, les cortex moteurs M1 ne semblent pas être la structure responsable

---

3. Par exemple, les activités de basse fréquence de type VCN (variation contingente négative) pourraient se sommer aux activités d'intérêt

de la sélection de la réponse telle que proposée dans les modèles sans étape d'association S-R. En revanche, nous avons envisagé qu'ils soient le lieu de la mise en place d'un système de prévention de l'erreur. Une élévation du seuil de réponse pourrait être implémentée suite à la détection d'un risque d'erreur.

Par ailleurs, la sensibilité à la compatibilité de la N-40 et de l'activation couplée au décalage temporel entre les deux activités suggère une relation hiérarchique entre-elles. La N-40 refléterait l'étape de sélection de la réponse et plus précisément, pourrait être impliquée dans le rappel des règles d'association S-R. L'activité de la N-40 se répercuterait sur l'activation du cortex M1 controlatéral, ce qui expliquerait l'augmentation de l'amplitude de l'activation. En revanche, nous ne sommes pas en mesure de savoir si l'inhibition est pilotée par la N-40.

Enfin, les composants EEG des deux études diffèrent uniquement au niveau de la négativité enregistrée sur le cortex prémoteur ipsilatéral à la réponse. Celle-ci est sensible à la compatibilité dans la tâche de Simon, et pas dans la tâche d'Eriksen. Selon la taxonomie décrite par Kornblum *et al.* (1990), ces deux tâches feraient appel à des formes de conflit différentes, cette différence entre les deux tâches pourrait expliquer les résultats.



## Chapitre 6

# Modulations attentionnelles des aires perceptives induites par le contexte de la tâche : Étude IRMf

### 6.1 Introduction

Afin de rester adapté à l'environnement, les êtres vivants les plus évolués ont la faculté d'ajuster leurs comportements et de prendre de nouvelles décisions lorsque la situation dans laquelle ils se trouvent l'exige. Nous nous intéressons ici aux mécanismes neuronaux sous-tendant la mise en place de ces ajustements comportementaux. En laboratoire, ces mécanismes de flexibilité peuvent être étudiés dans des tâches de TR de choix, en étudiant sur les effets séquentiels, c'est à dire en analysant la performance à l'essai  $n$  en fonction de la nature de l'essais  $n-1$ . Par exemple, après une erreur, on observe un allongement du temps de réaction ainsi qu'une baisse de la probabilité de commettre une nouvelle erreur. Des ajustements comportementaux sont également mis en évidence lorsque l'on manipule la compatibilité des stimuli. Par exemple dans la tâche de Simon, les sujets doivent répondre en fonction de la couleur d'un stimulus. Ce stimulus est présenté à gauche ou à droite d'un point de fixation. On définit comme "compatibles" les essais pour lesquels la réponse requise (gauche ou droite) coïncide avec le côté de présentation du stimulus. À l'inverse, on définit comme "incompatibles" les essais pour lesquels la réponse requise ne coïncide pas avec le côté de présentation du stimulus. En moyenne, les sujets sont plus rapides pour les stimuli compatibles.

L'effet de compatibilité correspond à la différence moyenne des TR des essais incompatibles et compatibles (Craft et Simon, 1970; Hedge et Marsh, 1975). Cet effet de compatibilité est modulé par la nature du stimulus à l'essai précédent. Suite à un essai compatible, l'effet de compatibilité est important, alors qu'il est largement réduit suite à un essai incompatible (Gratton *et al.*, 1992; Stürmer *et al.*, 2002). Ces effets séquentiels ont été interprétés comme étant le fruit d'un ajustement comportemental. Quels sont les mécanismes physiologiques pouvant rendre compte de ces effets comportementaux ?

Il a été proposé que les ajustements comportementaux qui se mettent en place à la fin de l'essai soient le fruit d'une modulation attentionnelle pilotée par le cortex préfrontal. Un biais perceptif serait implémenté au niveau des aires chargées du traitement du stimulus. Cette idée est explicite dans le modèle du Conflit (Botvinick *et al.*, 2001; Yeung *et al.*, 2004) ainsi que dans le modèle ERN-RL (Holroyd et Coles, 2002; Holroyd *et al.*, 2005). Dans ces deux modèles, le CCA émet un signal indiquant la nécessité de refocaliser l'attention sur les caractéristiques pertinentes de la tâche. Ce signal est transmis aux processus attentionnels (cortex préfrontal) qui organise le biais perceptif au niveau de la couche d'entrée (aires perceptives visuelles) en augmentant le niveau de base des unités de la couche d'entrée chargées de traiter la cible, et en diminuant le niveau de base des unités chargées du traitement des distracteurs. Si les hypothèses issues de ces deux modèles concernant la nature des signaux indiquant la nécessité d'une refocalisation attentionnelle ont été mis en difficulté, l'hypothèse de refocalisation attentionnelle concernant la mise en place des ajustements comportementaux reste théoriquement valide. Un moyen de tester cette hypothèse est de chercher des marqueurs physiologiques de ces modulations attentionnelles. Or, en IRMf il a été montré que le niveau d'attention porté aux paramètres des stimuli module l'activité des zones impliquées dans les traitements de ces paramètres (Corbetta *et al.*, 1990; Chawla *et al.*, 1999; Coull *et al.*, 2004; O'Craven *et al.*, 1997). Si les ajustements comportementaux sont bien le fruit d'une modulation attentionnelle des aires chargées du traitement des stimuli, alors nous devrions observer des modulations d'activité dans ces aires cérébrales suite aux essais incompatibles corrects. Cette modulation attentionnelle peut concerner le traitement de l'attribut pertinent et/ou le traitement de l'attribut non-pertinent. Les modèles du conflit et ERN-RL ont proposé que la modulation attentionnelle suscite conjointement une augmentation du niveau de base de l'aire responsable du traitement de l'attribut pertinent ainsi qu'une diminution du niveau de base du traitement de l'attribut non-pertinent. Selon ces deux modèles, nous devrions observer après un essai incompatible correct, une augmentation d'activité dans la zone chargée du traitement de l'attribut pertinent et une baisse dans celle chargée du traitement de l'attribut non-pertinent.

Cette hypothèse a été testée par Egner et Hirsch (2005a). Ils ont utilisé une tâche adaptée de la tâche de Stroop (figure 2.9 page 57). Les stimuli étaient composés d'un visage et d'un nom de personnages célèbres (hommes politiques ou acteurs) écrit en toutes lettres. Les sujets réalisaient deux tâches. Dans chacune de ces tâches, ils devaient répondre en fonction de la catégorie (hommes politiques ou acteurs), soit du visage (tâche 1) soit du nom (tâche 2). Lorsque le visage et le nom appartenaient à une même catégorie, l'essai était dit "compatible". Au contraire, si le nom et le visage appartenaient à des catégories différentes, l'essai était dit "incompatible". Les auteurs se sont intéressés aux variations du niveau d'activité dans l'aire spécifique au traitement des visages (*fusiform face area* - FFA). Celle-ci est impliquée dans le traitement de l'attribut pertinent dans la tâche de catégorisation des visages, et dans le traitement de l'attribut non-pertinent dans la tâche de catégorisation des noms.

Dans la tâche de catégorisation sur les noms, aucune modulation de la FFA n'est observée en fonction de la nature de l'essai en cours et la nature de l'essai précédent. Dans la tâche de catégorisation sur les visages, la FFA est plus activée dans les essais incompatibles lorsqu'ils sont précédés d'un essai incompatible ( $II > CI$ )<sup>1</sup>. Les auteurs argumentent que la quantité de traitement alloué à l'analyse des visages dans les essais incompatibles est plus importante lorsque l'essai précédent est incompatible plutôt que compatible. Ce qui suggère que le sujet re-focalise son attention sur le visage après un essai incompatible. Néanmoins, pour que l'hypothèse de la refocalisation attentionnelle soit suffisante pour expliquer ces résultats, il aurait fallu que l'activité de la FFA ne soit dépendante que de la nature de l'essai précédent. Or, l'activité de la FFA interagit avec la compatibilité à l'essai en cours et la compatibilité à l'essai précédent : la FFA est plus activée dans les essais compatibles que dans les essais incompatibles lorsque le stimulus précédent est compatible ( $CC > CI$ ). Les auteurs ont vraisemblablement observé un mécanisme physiologique prenant place en cours d'essai qu'ils ne parviennent pas à expliquer. Ces résultats suggèrent qu'il existe une modulation attentionnelle des aires chargées du traitement des caractéristiques pertinentes de la tâche en fonction de la nature de l'essai précédent et de la nature de l'essai en cours. Quoiqu'il en soit, il apparaît difficile de dissocier la part d'augmentation d'activité de l'aire perceptive due à un ajustement comportemental entrepris entre deux essais, de celle due au traitement du stimulus en cours.

Nous avons cherché à clarifier cette question en essayant d'isoler les mécanismes prenant place

---

1. Soit CC les essais compatibles précédés d'un essai compatible, CI les essais incompatibles précédés d'un essai compatible, IC les essais compatibles précédés d'un essai incompatible et II les essais incompatibles précédés d'un essai incompatible.

dans l'intervalle inter-essais dans une étude IRMf. L'objectif de ce travail étant de comparer, dans l'intervalle inter-essais, l'activité des aires chargées du traitement des caractéristiques du stimulus, il fallait être capable (1) de décorrélérer les activités induites par les traitements du stimulus en cours de celles induites par les traitements du stimulus précédent, et (2) de dissocier anatomiquement les aires perceptives en charge du traitement de l'attribut pertinent et de l'attribut non-pertinent.

Nous avons donc construit une tâche répondant à ces critères. Pour cela, nous avons choisi d'utiliser la couleur (préférentiellement traitée dans l'aire visuelle V4, Livingstone et Hubel, 1988; Zeki *et al.*, 1991) comme attribut pertinent et le mouvement (préférentiellement traitée dans l'aire visuelle hMT/V5, Watson *et al.*, 1993; Tootell *et al.*, 1995) comme attribut non-pertinent. Chawla *et al.* (1999) ont en effet montré qu'une focalisation attentionnelle augmentait le niveau d'activité de base des régions V4 et hMT/V5 lorsque les sujets devaient prêter attention respectivement à la couleur ou au mouvement.

Classiquement dans les tâches de TR, les essais sont séparés par un intervalle inter-essais court. Si ce rythme est adapté à l'EEG, il ne permet pas en IRMf de dissocier l'activité due au traitement du stimulus de l'activité due à la mise en place de la refocalisation attentionnelle<sup>2</sup>. En revanche, lorsque l'intervalle inter-essais est trop long, on peut craindre que l'influence de l'essai précédent diminue avec l'augmentation de l'intervalle inter-essais. Afin d'éviter les deux problèmes, nous avons choisi de présenter les stimuli par paires et d'estimer la réponse hémodynamique (BOLD - blood oxygenation level-dependant, Ogawa *et al.*, 1990) calée sur le premier stimulus de la paire. Dans notre design, l'intervalle entre les deux essais de chaque paire est court (entre 500 et 1500 ms), alors que l'intervalle entre les paires est long (entre 3000 et 15000 ms) de façon à ce que le premier essai de la paire suivante ne soit pas influencé par le second essai de la paire précédente.

## 6.2 Méthode

### 6.2.1 Les sujets

Quinze sujets hommes volontaires ont participé à cette expérience. Ils étaient âgés de 20 à 41 ans, l'âge moyen étant de 25.6 ans. Chacun a signé un formulaire de consentement de participation ainsi qu'une notice d'information destinée aux volontaires sains. Le protocole a reçu l'accord du C.C.P.P.R.B. Marseille 1 sous le numéro CM1 05/40.

---

2. C'est précisément le problème rencontré par Egner et Hirsch (2005a) en IRMf.

### 6.2.2 Les stimuli

Les stimuli visuels, conçus par nous même, étaient composés de losanges de couleur bleu ou verte, qui se déplaçaient uniformément de gauche à droite (ou de droite à gauche) sur un fond noir. L'impression donnée par le stimulus était de regarder des losanges de couleur se déplacer à travers une fenêtre fixe. Ces stimuli étaient projetés au centre de l'écran.

Les sujets devaient répondre par un appui sur un bouton poussoir placé dans leur main gauche ou dans leur main droite, en fonction de la couleur du stimulus ; la direction du mouvement n'était donc pas pertinente. Pour 7 des 15 sujets, la consigne était d'appuyer sur le bouton placé dans la main gauche lorsque le stimulus était vert, et sur le bouton droit lorsque le stimulus était bleu. Les autres sujets avaient la consigne inverse. On demandait également aux sujets de répondre le plus vite possible en évitant de commettre des erreurs. La couleur représentait donc l'attribut pertinent et le mouvement l'attribut non-pertinent. Ces deux attributs ont été choisis pour pouvoir dissocier anatomiquement les régions chargées de leur traitement.

Les essais compatibles sont ceux pour lesquels les losanges se dirigent du côté de la réponse correcte. Les essais incompatibles sont ceux pour lesquels les losanges se dirigent du côté de la réponse incorrecte.

### 6.2.3 Dessin expérimental

Puisque nous étions intéressés par les activités dans l'intervalle inter-essais, nous ne pouvions pas utiliser un dessin expérimental événementiel classique du fait des contraintes inhérentes à l'IRMf. En effet, décorréler les activités induites par le traitement de chaque essai nécessite l'utilisation d'intervalles inter-essais très longs et très variables (entre 3 et 15 secondes). Avec des intervalles inter-essais aussi longs, il n'était pas garanti de retrouver les effets séquentiels décrits plus haut. Nous avons donc choisi de présenter les stimuli par paires, avec des intervalles inter-essais courts et des intervalles inter-paires longs et variables. Nous pouvions ainsi décorréler les activités dues à la présentation des différentes paires.

La tâche se déroulait de la façon suivante (figure 6.1). Une croix de fixation était présentée au centre de l'écran pendant 750 ms indiquant au sujet l'apparition prochaine du stimulus. Après la présentation de cette croix, le premier stimulus de la paire (stimulus 1) était présenté pendant 750 ms. Le sujet devait répondre le plus vite possible en fonction de la couleur des losanges. La disparition du stimulus était immédiatement suivie de l'apparition d'une croix de fixation qui restait à l'écran entre 500 et 1500 ms (intervalle inter-essais suivant une distribution uniforme).

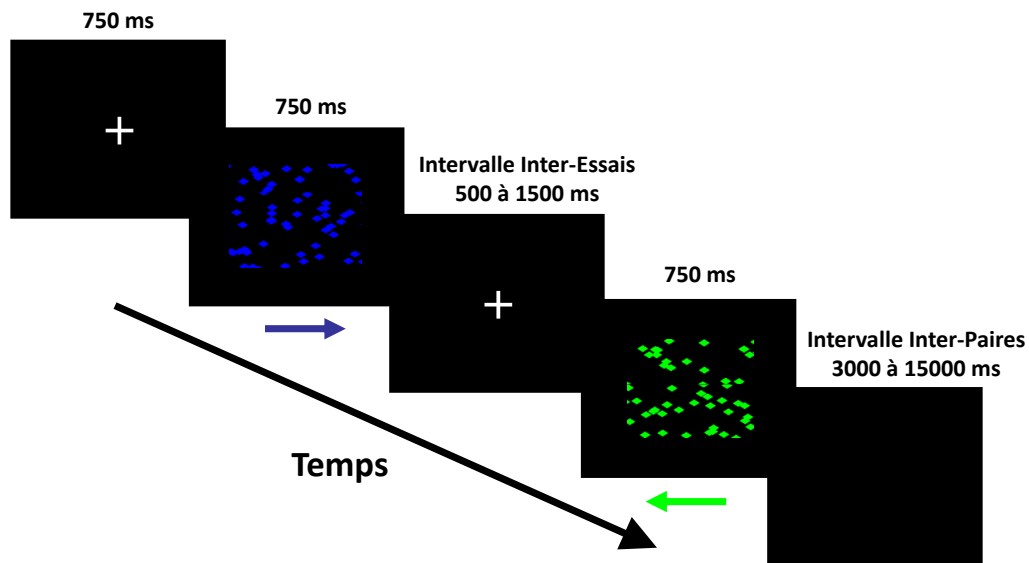


FIGURE 6.1 – Design expérimental

Ensuite le second stimulus est présenté pendant 750 ms. Le sujet devait à nouveau répondre le plus vite possible. Enfin l'écran restait noir pendant une durée comprise entre 3 et 15 s (intervalle inter-paires, suivant une distribution exponentielle décroissante de moyenne 6 secondes, de minimum 3 s et de maximum 15 s).

Les sujets effectuaient 4 sessions de 64 paires de stimuli présentés dans un ordre pseudo-aléatoire. Les stimuli étaient projetés sur un écran grâce à un vidéoprojecteur situé dans la chambre du scanner. Le temps d'exposition des stimuli est contrôlé par un système faisant intervenir une diode photosensible dont la résolution temporelle est supérieure à celle du vidéoprojecteur. Le sujet, à l'intérieur du scanner, percevait les stimuli par l'intermédiaire d'un miroir placé au dessus de ses yeux et orienté vers l'écran.

#### 6.2.4 Acquisition des données

L'expérience a lieu au centre d'IRM fonctionnelle de Marseille. Les données ont été acquises sur un imageur corps-entier 3-Teslas MEDSPEC 30/80 ADVANCE (Brücker, Ettlingen, Allemagne), équipé d'une antenne tête circulaire polarisée. Dans un premier temps, une séquence de repérage très rapide permettant d'identifier la commissure antérieure et la commissure postérieure de façon à choisir le plan de coupe était réalisée. Pour l'imagerie fonctionnelle, nous avons utilisé une séquence echoplanar (EPI) pondérée en T2\*, couvrant tout le cerveau avec 36 coupes axiales entrelacées de

3.5 mm d'épaisseur (temps de répétition = 2400 ms, temps d'écho = 35 ms, flip angle = 80°, field of view = 19.2 × 19.2 cm, matrice de voxels de 3 × 3 × 3 mm de 64 × 64). 180 volumes étaient acquis dans chaque session. A la fin de la session expérimentale, une image structurale haute résolution pondérée en T1 (séquence inversion-recovery, 1 × 0.75 × 1.22 mm) couvrant tout le cerveau et parallèle au plan Commissure Antérieure-Commissure postérieure défini à la première étape était acquise.

### 6.2.5 Traitement des données

Les données ont été traitées avec le logiciel SPM5 (<http://www.fil.ion.ucl.ac.uk/spm/>).

#### Prétraitements

En amont des analyses statistiques, nous devons effectuer une série de prétraitements listés ci-dessous.

Dans un premier temps, le décalage d'acquisition entre les coupes au sein d'un même volume a été corrigé. En effet, les différentes coupes d'un même volume ne sont pas acquises simultanément, mais successivement en mode entrelacé pendant une durée égale au temps de répétition (TR=2.4 s). L'étape de **recalage temporel** ("*slice timing*") consiste à ramener, par interpolation temporelle, l'instant d'acquisition de toutes les coupes à un instant commun. Au moment des traitements statistiques, on considérera que toutes les coupes du volume ont été acquises simultanément.

L'étape suivante est l'étape de **réalignement** ("*realignment*") qui corrige les mouvements de tête. Cette étape consiste à réaligner toutes les images acquises au cours de la session expérimentale sur la première image enregistrée en utilisant les informations sur les mouvements de tête effectués par le sujet au cours des différentes sessions.

La troisième étape est l'étape de **normalisation** ("*normalize*") qui consiste à déformer les images par référence au modèle ("*template*") EPI, fourni avec SPM5.

La quatrième étape est l'étape de **lissage spatial** ("*smooth*") qui applique un filtre gaussien spatial de 8 mm FWHM (full width at half maximum).

#### Analyses statistiques

L'analyse statistique était basée sur le modèle linéaire général (Friston *et al.*, 1995) dont l'équation est la suivante :  $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ .

$\mathbf{Y}$  est la matrice des données (valeur du signal BOLD mesurée pour chaque échantillon de temps et pour chaque voxel) ;

$\mathbf{X}$  est la matrice de design expérimental (décours temporels des stimulations associées à chacune des composantes cognitives élémentaires (régresseurs), convolués par un modèle de réponse hémodynamique fixe et connue : la fonction de réponse hémodynamique (HRF)) ;

$\beta$  est la matrice des paramètres de contribution de chaque composante cognitive au sein du signal de chaque voxel (paramètre à estimer) ;

$\varepsilon$  est la matrice modélisant le bruit.

Ce type d'analyse statistiques permet de rechercher les zones dont le signal BOLD covarie de manière significative avec le décours temporel de la stimulation c'est à dire d'estimer les contributions  $\beta$  de chaque condition expérimentale au signal recueilli. Les comparaisons planifiées, ou contrastes, sont des tests statistiques sur des combinaisons linéaires des paramètres estimés du modèle ( $\beta$ ).

Pour nos analyses, nous avons construit deux régresseurs. Le premier régresseur contenait les latences de présentation du premier stimulus des paires dont le premier stimulus était compatible (CC et CI). Le deuxième régresseur contenait les latences de présentation du premier stimulus des paires dont le premier stimulus était incompatible (IC et II). Autrement dit, nous avons distingué les paires dont le premier stimulus était compatible de celles dont le premier stimulus était incompatible. Ces régresseurs ont été convolués avec la fonction de réponse hémodynamique canonique proposée par SPM5. Les coefficients  $\beta$  ont été estimés pour chaque voxel. Une fois les paramètres du modèle estimés, nous avons calculé deux contrastes statistiques pour chaque sujet "compatible - incompatible" et "incompatible - compatible".

La seconde étape a consisté en une analyse en effets aléatoires (*"random effects"*). Cette analyse correspond à une analyse de groupe prenant en compte la variance inter-individuelle. Il s'agit d'effectuer un test  $t$  sur les images de contrastes de tous les sujets. Cette analyse renseigne sur les régions qui sont activées par les mêmes conditions à travers les sujets.

## 6.3 Résultats

### 6.3.1 Résultats Comportementaux

Les analyses ont été effectuées sur les paires pour lesquelles les réponses aux deux essais étaient correctes. Nous avons analysé les Temps de Réaction (TR) du 1<sup>er</sup> essai et du 2<sup>e</sup> essai de la paire



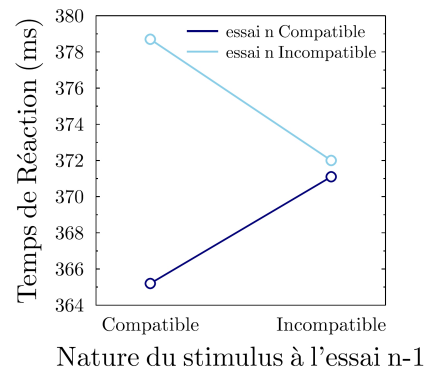


FIGURE 6.2 – **Effets séquentiels sur les TR du 2<sup>e</sup> essai de la paire.** L’effet de compatibilité sur le TR est important lorsque l’essai précédent (essai n-1) est compatible, et largement réduit lorsque l’essai précédent est incompatible (14 *versus* 1 ms).

en fonction de la compatibilité du stimulus. Nous n’observons pas d’effet de compatibilité sur les TR du 1<sup>er</sup> essai de la paire ( $F_{1,14} = 0.02, p = 0.88$ ). Le TR moyen pour un essai compatible est de 379.3 ms, et pour un essai incompatible est de 379.5 ms. Ce résultat suggère que le traitement du 1<sup>er</sup> stimulus de la paire est effectué de manière équivalente que le stimulus soit compatible ou incompatible.

L’effet de compatibilité sur le TR du 2<sup>e</sup> essai de la paire est significatif ( $F_{1,14} = 10.1, p < 0.01$ ). Les TR sont en moyenne plus courts dans les essais compatibles que dans les essais incompatibles (368 ms *versus* 375 ms). Cet effet de compatibilité au 2<sup>e</sup> essai est modulé par la nature du stimulus présenté au 1<sup>er</sup> essai ( $F_{1,14} = 17.9, p < 0.001$ ). La taille moyenne de l’effet de compatibilité est largement réduite après un essai incompatible (1 ms *versus* 14 ms après un essai compatible, voir figure 6.2).

### 6.3.2 Résultats IRMf

Pour l’analyse des données IRMf, nous avons modélisé les paires d’essais uniquement sur la base de la nature du 1<sup>er</sup> essai quelle que soit la nature du 2<sup>e</sup> essai. Deux régresseurs ont été utilisés, l’un pour les paires commençant par un stimulus compatible et l’autre pour celles commençant par un stimulus incompatible (respectivement les régresseurs “Compatible” et “Incompatible”). L’objectif était de mesurer des différences d’activité dans les aires chargées du traitement du stimulus suite à la présentation du premier stimulus. Nous nous sommes donc intéressés aux contrastes “Incompatible - Compatible” et “Compatible - Incompatible”.

Tout d’abord, les activations dans les régions frontales observées dans le contraste “Incompatible

Localisations Anatomiques	Coordonnées de Talairach (mm)			<i>t</i>	<i>p</i>
	<i>x</i>	<i>y</i>	<i>z</i>		
Régions frontales					
Gyrus frontal médian (BA 32)	-3	6	42	3.34	<i>p</i> = .005
Gyrus frontal moyen droit (BA 9)	30	42	36	5.8	<i>p</i> = .001
Gyrus frontal moyen gauche (BA 8/9)	-27	27	42	4.17	<i>p</i> = .005
	-33	33	36	3.43	<i>p</i> = .005
Régions temporo-occipitales					
Lobe postérieur gauche	-36	-60	-27	3.79	<i>p</i> = .005

FIGURE 6.3 – Table des résultats des activations dans les régions frontales et temporo-occipitales observées dans le contraste “Incompatible - Compatible”.

- Compatible” font apparaitre le cortex cingulaire antérieur et des activités bilatérales préfrontales dorsolatérales (voir table des résultats figure 6.3 et les représentations des activations figure 6.4). Ces résultats sont cohérents avec ceux de la littérature rapportant des activités dans ces deux régions suite à un essai incompatible (Kerns *et al.*, 2004). La question scientifique portait sur d’éventuelles activités dans les régions temporo-occipitales dans les aires chargées de l’attribut pertinent (V4 pour le traitement de la couleur) et de l’attribut non-pertinent (hMT/V5 pour le traitement du mouvement). Les résultats du contraste “Incompatible - Compatible” ont révélé une activité significative dans le lobe postérieur gauche que nous attribuons à V4 (voir figure 6.6 de gauche). Les coordonnées Talairach sont proches de la région identifiée comme étant V4 par Coull *et al.* (2004). Nous répliquons les résultats d’Egner et Hirsch (2005a) montrant une modulation de l’aire chargée de l’attribut pertinent après un essai incompatible.

La question plus intéressante était de savoir si l’on observait une variation d’activité dans l’aire hMT/V5 dans le contraste “Compatible - Incompatible”. Les résultats ont révélé une augmentation d’activité dans le gyrus temporal inférieur/moyen droit qui correspond à hMT/V5 (voir table figure 6.5 et les figures 6.6). Ces résultats suggèrent qu’il y a davantage d’activité dans l’aire hMT/V5 après un essai compatible qu’après un essai incompatible ce qui indique une suppression de l’attribut non pertinent après un essai incompatible. Il convient de noter que nous observons également une aire dans le lobe postérieur gauche proche de V4. Aucune activité dans les régions frontales n’apparaît dans ce contraste (voir les “glass brain” de la figure 6.6 à droite). En revanche, nous avons observé des activités dans les régions temporo-occipitales (voir les “glass brain” de la figure 6.6 de gauche) .

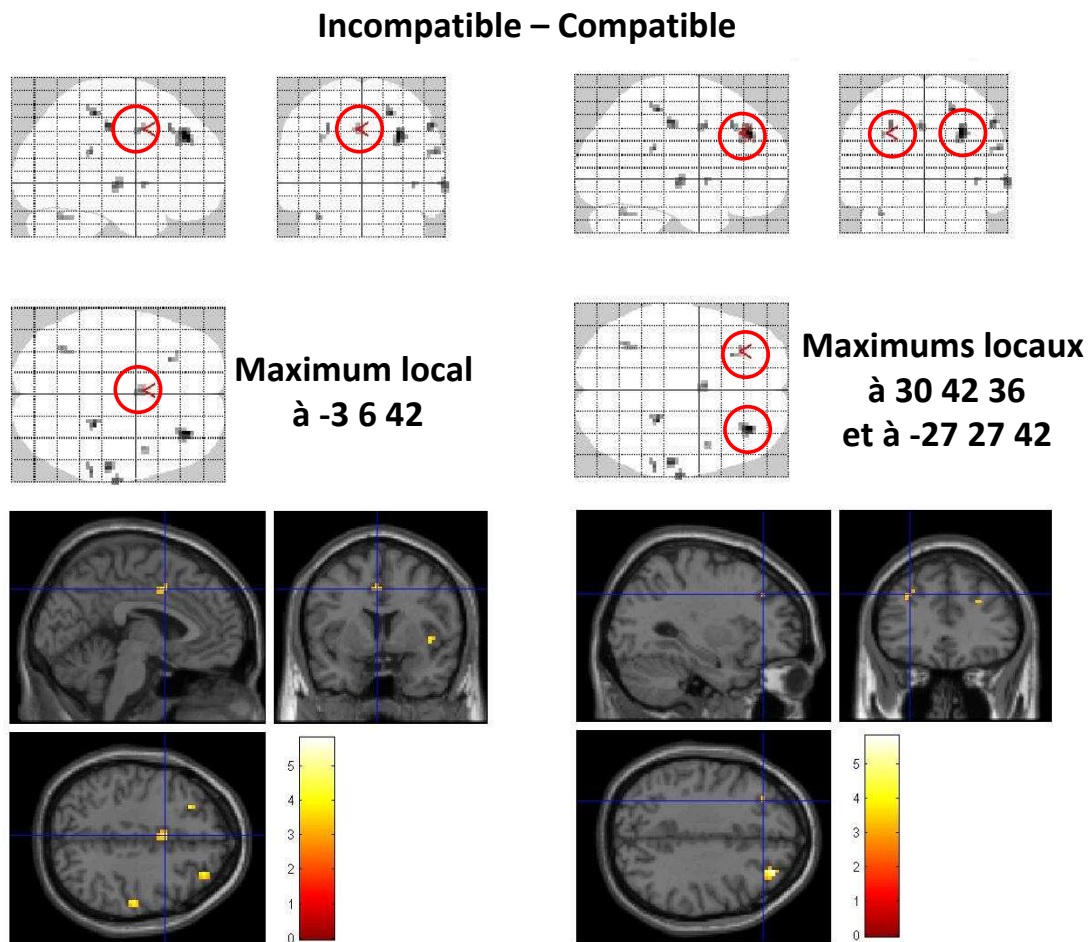


FIGURE 6.4 – Représentations anatomiques des régions frontales observées dans le contraste “Incompatible - Compatible”. Les deux graphiques du haut représentent les activités statistiquement significatives à  $p = .005$ , non corrigé, sous forme de glass brain (“le cerveau de verre de SPM”) que l’on utilise pour regarder globalement les aires activées en transparence selon trois points de vues. Les cercles rouges correspondent aux zones activées d’intérêt qui sont observées dans les deux illustrations du bas. Il s’agit des représentations des activations sur des coupes anatomiques dans les trois dimensions centrées sur l’activation dans le gyrus frontal médial (illustration de gauche) et sur le gyrus frontal moyen de l’hémisphère gauche et de l’hémisphère droit (illustration de droite).

Localisations Anatomiques	Coordonnées de Talairach (mm)			$t$	$p$
	$x$	$y$	$z$		
Régions temporo-occipitales					
Gyrus temporal inférieur droit	51	-54	-3	3.57	$p = .001$
Gyrus temporal moyen droit	45	-57	-3	3.42	$p = .001$
Lobe occipital gauche	-39	-75	-15	4.17	$p = .005$

FIGURE 6.5 – Table des résultats dans le contraste “Compatible - Incompatible”.

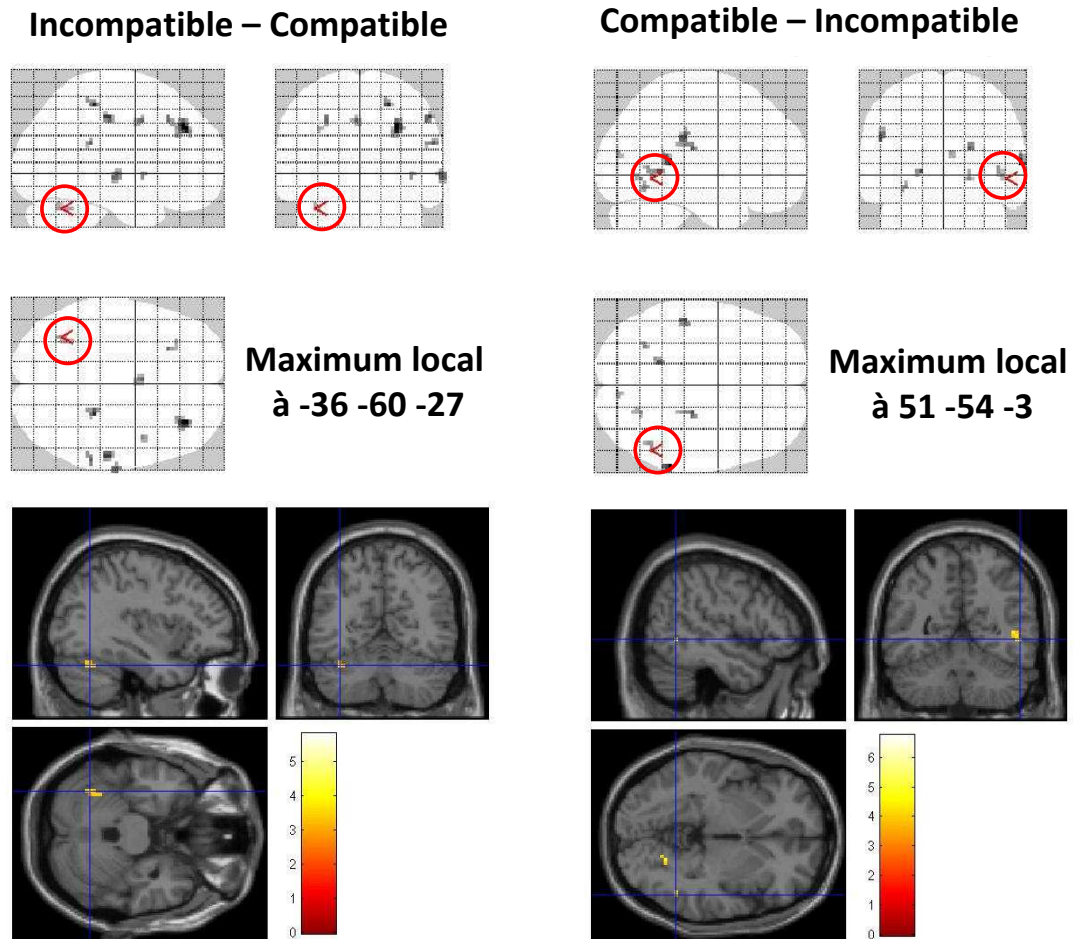


FIGURE 6.6 – Représentations anatomiques des régions temporo-occipitales observées dans les contrastes “Incompatible - Compatible” et “Compatible - Incompatible”. Les deux graphiques du haut représentent les activités statistiquement significatives à  $p = .005$ , non corrigé, sous forme de glass brain. Les cercles rouges correspondent aux zones activées d’intérêt qui sont observées dans les deux illustrations du bas. Il s’agit des représentations des activations sur des coupes anatomiques dans les trois dimensions centrées sur l’activation dans le lobe occipital gauche dans le contraste “Incompatible - Compatible” (illustration de gauche) et sur le gyrus temporal supérieur/moyen droit dans le contraste “Compatible - Incompatible” (illustration de droite).

## 6.4 Discussion

Les effets séquentiels observés par Gratton *et al.* (1992) dans la tâche d'Eriksen montrent qu'après un essai incompatible, la différence entre les TR compatibles et incompatibles est réduite. Cet effet, largement répliqué (Stürmer *et al.*, 2002; Burle *et al.*, 2005), a été interprété comme reflétant la refocalisation attentionnelle sur les caractéristiques pertinentes de la tâche suite à un essai incompatible (Gratton *et al.*, 1992; Botvinick *et al.*, 2001). Cependant, Mayr *et al.* (2003) ont remis en cause cette explication et ont proposé ces effets seraient dus à de simples effets d'alternance/répétition de stimuli. Il a été montré que la répétition d'un même stimulus diminue le TR (Cohen et Shoup, 1997; Hommel, 1998). La moitié des séquences d'essais CC et II bénéficient d'une répétition totale des caractéristiques du stimulus (par exemple dans la tâche d'Eriksen : HHH suivi de HHH et SSS suivi de SSS), pour l'autre moitié des séquences d'essais, les caractéristiques des stimuli sont totalement alternées (SSS suivi de HHH et HHH suivi de SSS). En revanche, les séquences d'essais CI et IC présentent une alternance et une répétition partielle du stimulus. En effet, soit les cibles sont identiques et les distracteurs différents, ou inversement les cibles sont différentes et les distracteurs identiques (par exemple pour les séquences IC, HSH suivi de SSS ou de HHH). Ainsi, la diminution des TR dans les essais CC et II par rapport aux séquences d'essais IC et CI, pourrait être imputable à la répétition du stimulus dans la moitié des séquences d'essais (ce dont ne bénéficient pas les séquences d'essais CI et IC). Pour Mayr *et al.* (2003), les différences de TR observées par Gratton *et al.* (1992) ne sont pas dues à la mise en place d'ajustements comportementaux mais plutôt dues aux effets d'alternance/répétition. Depuis, Ullsperger *et al.* (2005) ont montré que les effets séquentiels ne pouvaient pas seulement être dus à ces effets d'alternance/répétition. Bien qu'il existe de plus en plus d'arguments physiologiques capables d'expliquer les effets séquentiels observés sur les TR (implication du CCA et du DLPFC), le débat sur ce sujet ne semble pas tout à fait être tranché (voir Nieuwenhuis *et al.*, 2006).

Egner et Hirsch (2005a) ont testé directement l'hypothèse de la refocalisation attentionnelle comme mécanisme physiologique responsable des ajustements comportementaux. Ils ont cherché à observer des modulations des aires chargées du traitement des caractéristiques du stimulus dans une tâche manipulant la compatibilité Stimulus-Réponse. Les résultats ont montré une modulation de l'attribut pertinent mais pas de l'attribut non pertinent suite à un essai incompatible. Cependant, leur design expérimental n'était pas optimal. Dans les deux tâches, les stimuli étaient composés de deux objets différents : un visage et un nom (voir figure 2.9 page 57). Le sujet devait classer soit le visage, soit le nom dans une des deux catégories proposées (homme politiques ou acteur de

cinéma). Ils se sont intéressés aux variations d’une même aire visuelle (la FFA - spécialisée dans le traitement des visages) dans les deux tâches. Dans la tâche de catégorisation des visages, la FFA était impliquée dans le traitement de l’attribut pertinent du stimulus alors que dans la tâche de catégorisation du nom, la FFA était impliquée dans le traitement de l’attribut non-pertinent. Egner et Hirsch (2005a) ont donc comparé les variations d’une même aire visuelle dans deux tâches différentes qui ne sont certainement pas équivalentes. Intuitivement, il apparaît plus facile de faire abstraction de l’attribut pertinent dans une tâche où les deux dimensions n’appartiennent pas au même objet que dans des tâches où l’attribut non-pertinent est intrinsèquement lié à l’attribut pertinent (comme dans la tâche de Simon). Les résultats obtenus par Egner et Hirsch (2005a) ont montré une modulation de l’activité dans la FFA dans la tâche de catégorisation des visages mais pas dans la tâche de catégorisation des noms. Les auteurs ont interprété ces résultats comme indiquant qu’en réponse à un grand conflit, les mécanismes de contrôle cognitif augmentent de façon transitoire l’activité dans les aires corticales chargées du traitement des informations pertinentes plutôt que d’inhiber l’activité dans les aires corticales chargées du traitement des informations non-pertinentes.

Nous avons cherché à répliquer ces résultats en cherchant à isoler l’activité métabolique entre deux essais. La refocalisation attentionnelle sur les caractéristiques pertinentes de la tâche, expliquant les ajustements comportementaux dans le modèle du conflit (Botvinick *et al.*, 2001), est attendue après un essai incompatible. Pour valider cette hypothèse, nous devons réussir à identifier les variations du niveau d’activité dans des aires visuelles suite à la présentation d’un essai compatible et d’un essai incompatible dans une même tâche, et qui ne seraient pas dues au traitement du stimulus précédent. Nous avons conçu une tâche sur le modèle de la tâche de Simon (Craft et Simon, 1970). Les sujets devaient répondre en fonction de la couleur de stimuli (préférentiellement traitée dans V4). L’interférence était créée par le mouvement qui animait les stimuli (préférentiellement traité dans hMT/V5). De cette façon, les attributs pertinent et non-pertinent des stimuli étaient intrinsèquement liés (*“la couleur bouge”*). Par ailleurs, les stimuli étaient présentés par paires pour que l’on puisse être capable d’analyser les différences d’activités des aires visuelles dans l’intervalle inter-essais.

Les temps de réaction (TR) des essais corrects en réponse au premier essai de la paire ne sont pas sensibles à la nature du stimulus. En revanche, ceux en réponse au deuxième essai de la paire sont sensibles à la nature du stimulus en cours (effet de compatibilité), ainsi qu’à la nature du stimulus précédent (effets séquentiels). Les TR sont en moyenne plus courts lorsque l’essai est compatible. Cet effet de compatibilité est important lorsque l’essai précédent est compatible, et est largement

réduit lorsque l'essai précédent est incompatible. Ces résultats comportementaux suggèrent que le traitement du premier essai de la paire n'est pas influencé par les paires précédentes, ni par la nature de l'essai en cours. En revanche, les TR du deuxième essai de la paire montrent que le traitement de l'information dépend clairement de la nature du premier et du deuxième essai. Nous retrouvons les résultats comportementaux rapportés dans beaucoup d'études (Gratton *et al.*, 1992; Stürmer *et al.*, 2002; Burle *et al.*, 2005).

Les résultats IRMf montrent que dans l'intervalle inter-essais après un essai incompatible, une activation du réseau classiquement rapporté comme étant impliqué dans le contrôle cognitif. Il s'agit du CCA (Pardo *et al.*, 1990; Carter *et al.*, 1998, 2000; Pochon *et al.*, 2008) et du cortex préfrontal dorsolatéral (MacDonald *et al.*, 2000; Kerns *et al.*, 2004; Egner et Hirsch, 2005b; Kerns, 2006).

Au niveau des aires chargées du traitement du stimulus, nous avons observé une plus grande activité dans V4 (aire en charge du traitement de l'attribut pertinent) après un essai incompatible. Nous confirmons le résultat obtenu par (Egner et Hirsch, 2005a) montrant une modulation de l'aire chargée de l'attribut pertinent après un essai incompatible. Ce qui est nouveau dans notre étude est que nous avons également observé une modulation de l'aire hMT/V5 en charge de l'attribut non-pertinent après un stimulus incompatible. Ce résultat est très intéressant, il suggère une diminution du niveau de base de l'activité chargée du traitement de l'attribut non-pertinent suite à un essai incompatible. Ce résultat peut être interprété de la façon suivante : après un essai compatible dans lequel le mouvement aidait à la sélection de la réponse, on devient plus laxiste et cela se traduirait par une tendance à utiliser l'information sur le mouvement pour répondre le plus vite possible. Sans chercher à favoriser particulièrement le mouvement dans le choix de la réponse, on attacherait plus d'importance à celui-ci. En revanche, après un essai incompatible on chercherait à diminuer l'influence de l'information du mouvement en abaissant le niveau d'activité dans la région hMT/V5 dans l'intervalle inter-essais. Ce résultat n'a pas été observé par Egner et Hirsch (2005a), mais il est en accord avec l'hypothèse de refocalisation attentionnelle avancée par le modèle du conflit et le modèle ERN-RL (Botvinick *et al.*, 2001; Holroyd *et al.*, 2005). Cette hypothèse propose la mise en place d'un double mécanisme intervenant entre deux essais lorsque c'est nécessaire. Il se traduit par une augmentation du niveau de base des aires chargées du traitement des caractéristiques du stimulus et qui diminue le niveau de base des aires chargées du traitement des caractéristiques non-pertinentes du stimulus.

La refocalisation attentionnelle sur les caractéristiques pertinentes de la tâche semble bien avoir lieu, et prendrait la forme d'une augmentation de l'activité de l'aire responsable du traitement des

caractéristiques du stimulus. Cette augmentation d'activité de l'attribut pertinent s'accompagne bien d'une diminution de l'activité responsable de l'attribut non-pertinent. Nous ajoutons que cette inhibition pourrait dépendre de la "dangerosité" de l'attribut non-pertinent, et donc serait spécifique à la situation dans laquelle se trouve le sujet. Ce qui pourrait expliquer pourquoi Egner et Hirsch (2005a) ne l'observent pas. Toutefois, le recrutement de cette refocalisation attentionnelle reste encore flou, et plusieurs points restent encore à éclaircir. De manière générale, beaucoup d'études placent le cortex préfrontal au centre de cette refocalisation attentionnelle (MacDonald *et al.*, 2000; Kerns *et al.*, 2004; Egner et Hirsch, 2005b; Kerns, 2006). Cette vision pose la question du recrutement du contrôle cognitif. Il a été proposé que le CCA, sur la base de la mesure du conflit entre les réponses, soit la structure qui sollicite le cortex préfrontal (Botvinick *et al.*, 2001; Yeung *et al.*, 2004). Cette hypothèse est aujourd'hui remise en cause (Burle *et al.*, 2005, 2008a; Carbonnell et Falkenstein, 2006; Masaki *et al.*, 2007; Nakamura *et al.*, 2005). D'autres études proposent que la CCA puisse être plus directement lié à la prise de décision (Cohen *et al.*, 2005; Rushworth *et al.*, 2007; Pochon *et al.*, 2008). Enfin, il a récemment été proposé que le CCA ait un rôle dans les processus d'anticipation au contrôle et donc dans la mise en place des ajustements comportementaux d'anticipation (Aarts *et al.*, 2008). Le rôle des structures frontales n'est pas encore bien déterminé.



Troisième partie

## Discussion Générale



## Chapitre 7

# Discussion Générale

### 7.1 Nature de la $N_E$

#### 7.1.1 $N_E$ et erreurs

La  $N_E$  a été initialement associée à la détection de l'erreur. Depuis, de nombreuses interprétations ont été proposées. Ce qui est commun à l'ensemble de ces propositions c'est que la  $N_E$  est investie d'un rôle important dans les processus de contrôle cognitif. Toutes les caractéristiques de cette onde laissent penser qu'elle reflète un mécanisme essentiel au sein des processus de contrôle de l'action. En effet, elle est temporellement très proche de la réponse et est modulée par la performance. Paradoxalement et malgré le très grand nombre d'études cherchant à préciser son rôle, celui-ci n'est toujours pas clairement identifié. Comme nous le verrons plus tard, la  $N_E$  n'est ni liée au ralentissement post-erreur, ni liée à la prise de conscience de l'erreur. Afin de préciser son rôle, nous essayerons d'abord d'en décrire les caractéristiques.

#### 7.1.2 $N_E$ et Conflit

Puisque les ébauches d'erreur déclenchent une  $N_E$ , elles peuvent être utilisées comme un marqueur du conflit : Burle *et al.* (2005) ont testé une hypothèse particulière du modèle du conflit. La détection d'un conflit à l'essai  $n$  déclenche un ajustement comportemental à l'essai  $n+1$ . Les auteurs ont montré qu'il n'existe pas de lien entre la présence du conflit et l'ajustement comportemental à l'essai suivant cette ébauche. Bien que les ébauches d'erreurs induisent une  $N_E$ , la modulation de l'effet de compatibilité ne varie pas en fonction de la présence ou non de l'ébauche d'erreur à l'essai précédent. Cette dernière n'est pas utilisée comme initiateur de l'ajustement comportemental.

De plus dans la première étude de ce manuscrit, nous sommes allés plus loin et nous avons quantifié le conflit sur la base des ébauches d'erreurs, puis nous avons comparé les résultats de la simulation du modèle du conflit avec les données expérimentales EEG. Les résultats de la cette étude montrent que le conflit ne peut pas rendre compte de la  $N_E$ . L'un des aspects intéressant de cette étude est que l'amplitude de la  $N_E$  de l'ébauche d'erreur est d'autant plus grande que la correction de cette ébauche arrive tardivement, ce qui suggère que cette  $N_E$  se développe jusqu'à ce que le "problème" soit réglé.

### 7.1.3 $N_E$ : un mécanisme général ?

Comme nous l'avons fait ressortir dans le chapitre sur les modèles (chapitre 2), aucun des modèles neuromimétiques ne peut rendre compte de la  $N_E$  dans les essais corrects. Tous considèrent, plus ou moins implicitement, que la  $N_E$  des essais corrects représente un autre mécanisme physiologique que celui qui déclenche la  $N_E$  des erreurs. Or les travaux exposés dans le chapitre 4 apportent des arguments forts en faveur de l'idée que la négativité dans les essais corrects est la manifestation du même processus physiologique que la  $N_E$  des erreurs. Vidal *et al.* (2000, 2003a) avaient déjà montré que la  $N_E$  des corrects avait la même latence et la même topographie. En utilisant l'analyse en composantes indépendantes, nous avons pu montrer que la  $N_E$  dans les trois types d'essais (corrects, ébauches d'erreur et erreurs) pouvait être expliquée par une seule et même composante indépendante. Nous avons également montré que les trois activités provenaient de la même source située dans la zone cingulaire rostrale.

Ces résultats nous amènent à repenser la complexité de cette onde. Elle semble être une onde très générale, présente après toute réponse, mais dont l'amplitude serait modulée par des facteurs liés à la performance et ou à la pathologie. Il reste à trouver le ou les facteur(s) modulateur(s) de cette  $N_E$ . Bien que rien de clair n'émerge, plusieurs candidats peuvent être suggérés.

Une autre aire frontale pourrait jouer un rôle modulateur de cette  $N_E$  (Swick *et al.*, 2002). Gehring et Knight (2002) ont montré qu'une lésion du cortex préfrontal altérerait la  $N_E$  des corrects. Son amplitude est augmentée, et atteint le niveau de celle des erreurs. Swick *et al.* (2002) ont montré des résultats similaires chez un patient avec une lésion de la partie rostrale du CCA, la  $N_E$  des corrects de ce patient était anormalement grande. Cette étude pose la question du générateur de la  $N_E$ . Puisque le CCA est une structure étendue, il est possible que la région lésée soit la partie rostrale du CCA ; la région caudale du CCA pourrait être le générateur de la  $N_E$ . Dans ce cas de figure, on pourrait supposer un effet modulateur d'une partie du CCA sur une autre qui serait le

générateur de la structure. Une autre possibilité est que la  $N_E$  serait générée par l'aire motrice supplémentaire et que la CCA soit l'un des modulateurs de cette structure.

Il existe des arguments en faveur d'un rôle des ganglions de la base dans les mécanismes du contrôle cognitif. Certaines théories supposent que la dopamine soit à l'origine de la genèse de la  $N_E$ . Cette hypothèse fait comme prédiction que la diminution de la DA devrait affecter l'amplitude de la  $N_E$  des erreurs et celle des corrects. Or Willemssen *et al.* (2009) ont montré que les patients parkinsoniens présentaient une diminution de l'amplitude de la  $N_E$  des erreurs et une augmentation de l'amplitude de la  $N_E$  dans les essais corrects.

Il serait envisageable que les processus par défaut sont de taille intermédiaire, et que cette taille par défaut serait modulée, soit positivement (augmentation de la  $N_E$  par exemple dans les erreurs), soit négativement (diminution de la  $N_E$  dans les corrects). Cette vision non standard nécessite néanmoins d'être mise à l'épreuve.

Au niveau fonctionnel, la  $N_E$  semble sensible à l'implication du sujet dans la tâche. Il semblerait que plus l'engagement des sujets est important, plus l'amplitude de la  $N_E$  dans les erreurs est importante (Gehring *et al.*, 1993). Les patients souffrants de TOC, connus pour leurs difficultés à se désengager d'une tâche, montrent une  $N_E$  de plus grande amplitude dans les erreurs et dans les corrects comparée aux sujets contrôles (Endrass *et al.*, 2008). Une étude intéressante cherchant à évaluer les effets de la fatigue mentale chez des sujets sains, montre que l'amplitude de la  $N_E$  diminue au fur et à mesure de la durée de la tâche (Boksem *et al.*, 2006). Néanmoins, si, au bout de deux heures de tâche, on apporte une nouvelle consigne au sujet qui permet de le motiver à nouveau, l'amplitude de la  $N_E$  est augmentée.

Nous constatons qu'aucune des propositions effectuées jusqu'à présent n'est capable de rendre compte des modulations de la négativité dans les trois types d'essais (erreurs, essais corrects et essais corrects avec ébauche d'erreur). Il n'est peut-être pas pertinent de chercher à rendre compte que l'activité du CCA par un seul mécanisme. Il est probable que le CCA soit une sorte de "centre intégrateur" vers lequel convergent beaucoup d'informations d'origine corticales et sous-corticales sur les informations perceptives, le contexte, l'état émotionnel du sujet (Quilodran *et al.*, 2008; Hajcak *et al.*, 2004). Cette vision du CCA pourrait expliquer l'hétérogénéité des résultats issus des études chez les patients.

## 7.2 Lien avec les autres ondes

### 7.2.1 $N_E$ et FRN

Dans une étude utilisant une méthodologie très similaire à celle utilisée dans le chapitre 4<sup>1</sup>, Gentsch *et al.* (2009) ont montré que la FRN semble être un composant qui soit dans la même classe que les négativités des erreurs et des essais corrects, rendant ce phénomène encore plus général. Pour valider complètement cette hypothèse, il faudrait analyser des situations expérimentales et vérifier que les facteurs modulant l'amplitude de la  $N_E$  affectent également l'amplitude de la FRN. Nous devrions également trouver des effets similaires sur la  $N_E$  et la FRN chez les patients.

### 7.2.2 $N_E$ et $N2$

La  $N2$  est une négativité enregistrée dans les essais corrects, dans l'amplitude de la  $N2$  est plus grande en situation incompatible que compatible. Elle a été rapportée calée sur le stimulus (Kopp *et al.*, 1996; Van Veen et Carter, 2002) et sur la réponse (Yeung *et al.*, 2004), et a été interprétée comme étant le reflet du conflit de réponse dans les essais corrects. Le modèle du conflit prédit que le conflit est présent avant la réponse dans les essais corrects, et après la réponse dans les erreurs. Ainsi Yeung *et al.* (2004) proposent que le corrélât électrophysiologique du conflit pré-réponse dans les essais corrects serait la  $N2$ , et celui du conflit post-réponse dans les erreurs serait la  $N_E$ .

L'enregistrement de l'activité EMG dans les muscles associés aux réponses permet d'observer des ébauches d'erreurs dans certains essais corrects (leur pourcentage est autour de 15% dans les tâches de TR). Une  $N_E$  au niveau central est observée au moment de l'ébauche d'erreur. Donc, si l'on ne distingue pas les essais corrects avec ébauche et les essais dits purs-corrects, on s'expose à mélanger les activités électrophysiologiques de ces deux types d'essais. Or la proportion d'ébauche d'erreurs étant plus importante en incompatible qu'en compatible, la proportion des essais corrects avec ébauche d'erreur inclus dans le moyennage est plus importante dans la condition incompatible que compatible. Si la  $N_E$  des ébauches peut émerger dans le grands moyennage de tous les essais corrects, alors il y a de fortes chances qu'elle soit de plus grande amplitude en condition incompatible, et pourrait expliquer, au moins en partie l'effet  $N2$  rapporté dans la littérature. Le problème se pose de la même façon pour les études IRMf : il est possible que l'activité du CCA rapportée dans les essais corrects corresponde à l'activité du CCA dans les ébauches d'erreur.

Une  $N2$  frontale n'a jamais été rapportée dans les données EEG après calcul du Laplacien

---

1. décomposition du signal EEG en sources indépendantes temporellement et avec des topographies fixes

de surface. Il n'est pas impossible que l'effet  $N2$  frontal calé sur le stimulus soit le reflet d'un recouvrement entre plusieurs activités. En utilisant la méthodologie des raster-plot<sup>2</sup>, Jung *et al.* (2001) montrent que la P3 suit clairement la réponse (voir leur figure 5B p.174). Puisque les TR dans les essais incompatibles sont plus longs que dans les essais compatibles, la latence de la P3 est plus tardive dans la condition incompatible que compatible. Si la latence de la  $N2$  est stable alors que la latence de la P3 varie, le recouvrement entre les activités n'est pas la même. Si l'on place deux dipôles dans des régions pariéto-occipales (l'un pour la  $N2$  et l'autre pour la P3) et que l'on fait varier la latence de la P3, on peut observer une augmentation de la  $N2$  lorsque la P3 est retardée, comme dans les essais incompatibles. La figure 7.1 montre le signal reconstruit en FCz. On observe très clairement une  $N2$  sur cette électrode, alors que celle-ci est générée dans la région pariéto-occipitale. En décalant la P3, on crée un "effet  $N2$ ", qui est maximal au niveau des électrodes frontales.

D'autre part, la  $N2$  observée par Yeung *et al.* (2004) dans les tracés EEG calés sur la réponse montrent qu'elle est de plus grande amplitude dans la condition incompatible que compatible. Cette activité ressemble beaucoup à la N-40 que nous avons étudiée au chapitre 5. Nous avons attribué un rôle de sélection de la réponse à cette activité. Comme pour la  $N2$  calée sur le stimulus, il est nécessaire de s'assurer que la différence d'amplitude de cette  $N2$  observée par Yeung *et al.* (2004) n'est pas due à la présence de la  $N_E$  déclenchée par les ébauches d'erreurs.

Concernant la N-40 observée dans les deux études au chapitre 5, nous ne pensons pas qu'elle puisse être due à la  $N_E$  déclenchée par certaines ébauches non détectables avec l'EMG pour la raison suivante : nous avons vu dans le premier chapitre expérimental que l'amplitude de la  $N_E$  est d'autant plus grande que le temps de correction est long. Or nous avons également montré que l'amplitude de la réponse musculaire est plus faible dans les essais où le temps de correction est court. Nous en déduisons que plus la réponse musculaire est petite, plus la  $N_E$  est de faible amplitude. Il devient alors difficile d'imaginer que la N-40 enregistrée dans les essais purs-corrects soit due à une  $N_E$  qui aurait été déclenchée par une activité de la réponse incorrecte suffisamment faible pour ne pas avoir été détectée sur les tracés EMG. En revanche, la  $N_E$  des ébauches pourrait participer à la  $N2$  de Yeung *et al.* (2004) puisque les auteurs n'ayant pas enregistré l'activité EMG, n'ont pas pu faire la distinction entre les essais purs-corrects et les essais corrects avec ébauche d'erreur.

De plus, nous supposons que la N-40 est générée par l'aire motrice supplémentaire et qu'elle reflète la difficulté de la tâche. Cette interprétation peut rendre compte de la différence dans les

---

2. Méthodologie que nous avons utilisée pour regarder les essais EEG "essais par essais" au chapitre 3

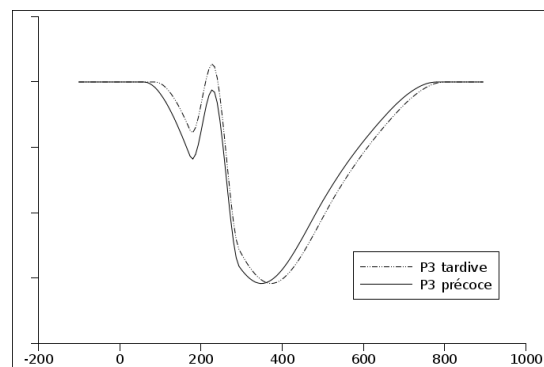
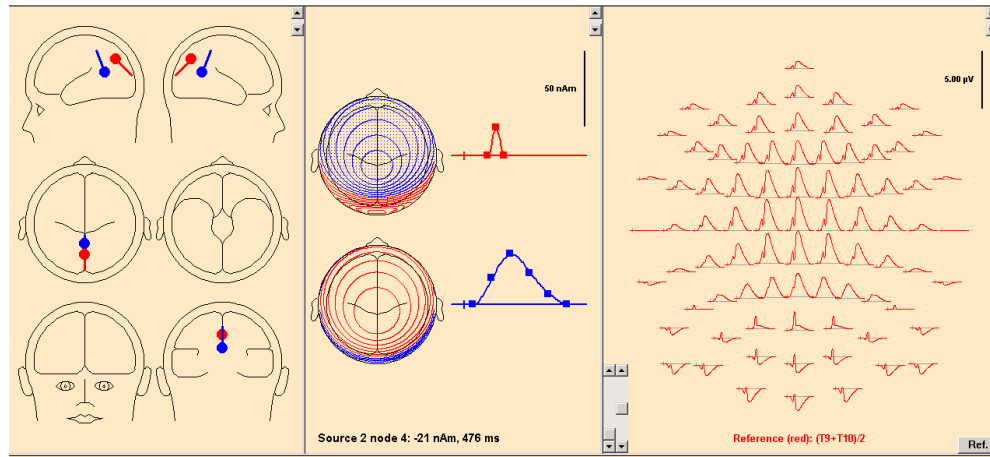


FIGURE 7.1 – **Simulation dipolaire de la N2 calée sur le stimulus et de la P3 calée sur la réponse.** Le dipôle rouge génère une activité autour de 200 ms post-stimulus dans les régions occipitales. Du fait de cette orientation, elle apparaît négative dans les régions frontales. Le dipôle bleu simule une activité de type P3 calée sur la réponse. Son orientation pointe vers les régions frontales, donc sa topographie est positive dans les régions frontales. Lorsque l'on fait varier la latence de la P3 (P3 précoce ou tardive), on observe un effet N2 dans les régions frontales (voir graphique du bas). La négativité autour de 200 ms apparaît plus grande dans la situation où la P3 est tardive. Cette simulation a été réalisée par Dipole Simulator (<http://www.besa.de/updates/tools/>).



essais compatibles et incompatibles, de sa diminution quand une information sur la nature de la réponse est donnée aux sujets avant l'apparition du stimulus (Carbonnell *et al.*, 2004), ainsi que son absence dans des tâches qui utilisent des association S-R plus simples (Burle *et al.*, 2004a; Meynier *et al.*, 2009).

## 7.3 Ajustements Comportementaux

### 7.3.1 Ralentissement post-erreur

Nieuwenhuis *et al.* (2001) montre que, pour les mêmes sujets, il semble exister un lien entre la  $P_E$  et le ralentissement post-erreur à travers la prise de conscience de l'erreur. La revue de littérature, résumée dans le tableau 1.17 (page 42) montre que ce lien très fort existe à travers les populations<sup>3</sup>.

Pour établir encore plus clairement ce lien, il faudrait reprendre la méthodologie de Gehring et Fencsik (2001) adaptée à la  $P_E$ . Les auteurs ont trié les essais qui suivent une erreur selon l'existence ou non d'un ralentissement post-erreur. On doit s'attendre à ce que la  $P_E$  soit plus grande dans les essais précédant un ralentissement post-erreur que dans les essais post-erreur sans ralentissement (pour lesquels elle pourrait même être absente).

Gehring et Fencsik (2001) ont trié de cette façon les essais et ont montré qu'il n'y avait pas de relation entre l'amplitude de la  $N_E$  et le ralentissement post-erreur (la  $N_E$  est toujours de même amplitude, que l'erreur soit suivie d'un ralentissement ou pas). La revue de littérature confirme la décorrélation entre l'amplitude de la  $N_E$  et le ralentissement post erreur<sup>4</sup>. Néanmoins, les patients lésés dans la zone du CCA ne montrent pas de ralentissement post-erreur, ce qui suggère qu'il existe tout de même un lien entre ralentissement post-erreur et le CCA (di Pellegrino *et al.*, 2007). Notons que si la zone lésée n'est pas le générateur de la  $N_E$  (Swick *et al.*, 2002), alors ces résultats ne sont pas contradictoires.

---

3. Sur les 5 études rapportant les données sur la  $P_E$  et le ralentissement post-erreur, toutes vont dans le sens d'une corrélation entre les deux éléments. 3 études rapportent que une  $P_E$  d'amplitude normale et la présence d'un ralentissement post-erreur, 2 études rapportent une  $P_E$  diminuée et un ralentissement post-erreur diminué voire absent

4. Sur les 4 études rapportant la présence du ralentissement post-erreur, 2 montrent une diminution de l'amplitude de la  $N_E$ . Sur les 3 études rapportant l'absence de ralentissement, 2 montrent une diminution de la  $N_E$  et 1 montre une augmentation de la  $N_E$ .

### 7.3.2 Accélération pré-erreur

Avant de commettre une erreur, les sujets ont tendance à accélérer (Laming, 1979). L'amplitude de la  $N_E$  dans les essais corrects est un indicateur du risque d'erreur à l'essai suivant. Ridderinkhof *et al.* (2003) ont étudié l'activité autour de la réponse correcte dans l'essai précédent l'erreur afin de chercher les signes avant-coureurs de l'erreur sur les données monopolaires. Les auteurs ont observé une augmentation de la positivité classiquement observée en EEG monopolaire après la réponse correcte dans les essais précédant une erreur. Après calcul du Laplacien de surface, Allain *et al.* (2004c) puis Cavanagh *et al.* (2009) ont montré qu'il s'agissait en fait d'une diminution de la  $N_E$  dans les essais corrects. Ce résultat pointe vers un lien entre l'activité du CCA (si la source de la  $N_E$  est bien le CCA) et la qualité du contrôle ou l'engagement du sujet dans la tâche.

### 7.3.3 Adaptation à la compatibilité

Egner et Hirsch (2005a) se sont intéressés à la mise en place de l'ajustement comportemental après un essai incompatible. Ils ont montré l'existence d'une augmentation d'activité dans l'aire chargée du traitement de l'attribut pertinent après un essai incompatible. En revanche, aucune différence d'activité n'a été observée dans l'aire chargée du traitement de l'attribut non-pertinent. Les modèles actuels prévoient plutôt à la fois une augmentation de l'importance accordée au traitement de l'attribut pertinent, et une diminution de la prise en compte de l'attribut non-pertinent. Dans notre étude IRMf, nous avons montré à la fois, une augmentation du traitement de l'attribut pertinent que nous interprétons comme une facilitation, et une diminution du traitement de l'attribut non-pertinent interprétée comme une inhibition.

### Qui déclenche l'ajustement comportemental ?

Selon plusieurs modèles, les régions préfrontales sont impliquées dans la mise en place de la refocalisation attentionnelle sur le traitement des caractéristiques pertinentes de la tâche. Plusieurs études ont observé, en IRMf, une activation du CCA et du cortex préfrontal dans les situations précédant une mise en place de l'ajustement comportemental (Kerns *et al.*, 2004; Kerns, 2006; Egner et Hirsch, 2005b). Le lien temporel entre l'activité du CCA et le cortex préfrontal a récemment été mis en évidence dans une étude EEG (Cavanagh *et al.*, 2009). Les auteurs ont montré une synchronie entre l'activité sur l'électrode placée au dessus du CCA et les électrodes placées au niveau du cortex préfrontal dorsolatéral au moment de la réponse.

Une autre théorie place le CCA au cœur de la mise en place des ajustements comportementaux.

Aarts *et al.* (2008) ont montré que le CCA était actif dans la période préparatoire lorsque l'indice indiquait au sujet si l'essai allait être incompatible. Le rôle du CCA a été proposé comme étant un rôle d'anticipation du contrôle. L'augmentation du CCA correspondrait à la mise en place de l'ajustement comportemental. Cette proposition est en ligne avec les travaux de di Pellegrino *et al.* (2007) qui montrent l'absence de modulation de l'effet de compatibilité après un essai correct incompatible et l'absence de ralentissement post-erreur chez les patients atteints de lésions cérébrales dans la zone cingulaire rostrale.

## 7.4 Conclusions

De ce travail de thèse ressortent trois points essentiels.

Tout d'abord, le phénomène physiologique sous-tendu par la " $N_E$ " reflète un processus générique, présent après toutes les réponses, et dont l'amplitude est modulée. Bien que la nature exacte des paramètres modulateurs reste à définir, cette nouvelle façon de concevoir ce signal électrophysiologique nous amène à repenser le questionnement sur cette onde : au lieu de se demander qu'elle est la nature des signaux qui la déclenche, il semble plus important de se demander quelle est la nature des signaux qui la module, tant d'un point de vue fonctionnel que physiologique.

Un deuxième point essentiel est que la régulation de l'interférence en fonction du contexte se fait bien, à la fois par une facilitation du traitement de l'information pertinente, et par l'inhibition des informations non pertinentes. Ces résultats remettent les phénomènes inhibiteurs au cœur des processus de contrôle, alors qu'ils avaient pu en être écartés suite aux travaux de Egner et Hirsch (2005a) : le commentaire que de Nieuwenhuis et Yeung (2005) font de cet article a un titre d'ailleurs évocateur : "*Losing our inhibition ?*". En montrant que le contrôle des informations se fait aussi à travers une inhibition, nos travaux renouvèlent le débat, et ouvrent de nouvelles perspectives.

Ce travail a également mis en évidence un réseau d'aires motrices et pré-frontales actives autour de la réponse. On voit en effet apparaître une séquence pré-SMA < cortex pré-moteur ipsilatéral < cortex moteur contralatéral < SMA/CCA ≤ pré-moteur contralatéral. Cette séquence, qui a pu être mise en évidence grâce à l'amélioration de la résolution temporelle induite par le calcul du Laplacien de surface, semble stable puisqu'elle se retrouve dans deux tâches différentes. Plusieurs de ces composantes présentent des sensibilités à des facteurs expérimentaux, tels que la compatibilité Stimulus-Réponse, ce qui pointe vers une implication des structures sous-jacentes dans les processus de décision et/ou d'évaluation. Comprendre le rôle précis de chacune de ces structures, en intégrant

leur dynamique temporelle, sera un des challenges des prochaines années.

Si les techniques d'imagerie non-invasives chez l'Homme ont permis des avancées considérables, certains aspects sont difficilement appréhendables chez l'Homme. Les travaux chez l'animal devraient permettre de tester plus spécifiquement certains points, notamment les hypothèses pharmacologiques qui sous-tendent la  $N_E$ . Pour cela, enregistrer l'équivalent chez le rat de ces signaux d'erreur représenterait une avancée considérable. Nous avons contribué au développement de tels enregistrements, en participant à la mise en place d'un poste d'enregistrement EEG chez le rat vigile, libre de ces mouvements (Roger *et al.*, 2009). L'équipe ayant, par ailleurs, déjà développé une modèle Rat de la tâche de Simon (Courtière *et al.*, 2007), tout est en place pour étudier ces phénomènes de contrôle chez l'animal, en couplant approches électrophysiologique et pharmacologique/lésionnelle.

# Bibliographie

- AARTS, E., ROELOFS, A. et van TURENNOUT, M. (2008). Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *Journal of Neuroscience*, 28(18):4671–4678.
- ADLEMAN, N. E., MENON, V., BLASEY, C. M., WHITE, C. D., WARSOFSKY, I. S., GLOVER, G. H. et REISS, A. L. (2002). A developmental fMRI study of the Stroop color-word task. *NeuroImage*, 16(1):61–75.
- ALLAIN, S. (2005). *Supervision de l'action et détection de l'erreur dans les activités sensori-motrices chez l'Homme : études comportementales et électrophysiologiques*. Thèse de doctorat, Université de Provence.
- ALLAIN, S., CARBONNELL, L., BURLE, B., HASBROUCQ, T. et VIDAL, F. (2004a). Electromyographic activities and the Ne. In ULLSPERGER, M. et FALKENSTEIN, M., éditeurs : *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring.*, pages 21–27, Leipzig. MPI of Cognitive Neuroscience.
- ALLAIN, S., CARBONNELL, L., BURLE, B., HASBROUCQ, T. et VIDAL, F. (2004b). Online executive control : an electromyographic study. *Psychophysiology*, 41:113–116.
- ALLAIN, S., CARBONNELL, L., FALKENSTEIN, M., BURLE, B. et VIDAL, F. (2004c). The modulation of the Ne-like wave on correct responses foreshadows errors. *Neuroscience Letters*, 372:161–166.
- AMIEZ, C., JOSEPH, J.-P. et PROCYK, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *European Journal of Neuroscience*, 21(12):3447–3452.
- AMIEZ, C., JOSEPH, J. P. et PROCYK, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cerebral Cortex*, 16(7):1040–1055.
- BABILONI, F., CINCOTTI, F., CARDUCCI, F., ROSSINI, P. M. et BABILONI, C. (2001). Spatial enhancement of EEG data by surface Laplacian estimation : the use of magnetic resonance imaging-based head models. *Clinical Neurophysiology*, 112:724–727.
- BARCH, D. M., BRAVER, T. S., SABB, F. W. et NOLL, D. C. (2000). Anterior cingulate and the monitoring of response conflict : evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience*, 12(2):298–309.
- BARTHOLOW, B. D., PEARSON, M. A., DICKTER, C. L., SHER, K. J., FABIANI, M. et GRATTON, G. (2005). Strategic control and medial frontal negativity : beyond errors and response conflict. *Psychophysiology*, 42(1):33–42.
- BARTO, A. G. (1995). Adaptive critics and the basal ganglia. In HOUK, J., DAVIS, J. et BEISER, D., éditeurs : *Models of information processing in the basal ganglia.*, pages 215–232. MIT Press, Cambridge.
- BATES, A. T., KIEHL, K. A., LAURENS, K. R. et LIDDLE, P. F. (2002). Error-related negativity and correct response negativity in schizophrenia. *Clinical Neurophysiology*, 113(9):1454–1463.
- BATES, A. T., LIDDLE, P. F., KIEHL, K. A. et NGAN, E. T. C. (2004). State dependent changes in error monitoring in schizophrenia. *Journal of Psychiatric Research*, 38(3):347–356.
- BATES, A. T., PATEL, T. P. et LIDDLE, P. F. (2005). External behavior monitoring mirrors internal behavior monitoring. error-related negativity for observed errors. *Journal of Psychophysiology*, 19(4):281–288.

- BELL, A. J. et SEJNOWSKI, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159.
- BERGER, B., GASPARD, P. et VERNEY, C. (1991). Dopaminergic innervation of the cerebral cortex : unexpected differences between rodents and primates. *Trends in Neurosciences*, 14(1):21–27.
- BERNSTEIN, P., SCHEFFERS, M. et COLES, M. (1995). “Where did I go wrong ?” A psychophysiological analysis of error detection. *Journal of Experimental Psychology : Human Perception and Performance*, 21:1312–1322.
- BLISS, T. V. et LØMO, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology*, 232(2):331–356.
- BOGACZ, R., BROWN, E., MOEHLIS, J., HOLMES, P. et COHEN, J. D. (2006). The physics of optimal decision making : a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4):700–765.
- BOKSEM, M. A. S., MEIJMAN, T. F. et LORIST, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biological Psychology*, 72(2):123–132.
- BOTVINICK, M. M., BRAVER, T. S., CARTER, C. S., BARCH, D. M. et COHEN, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108:624–642.
- BOTVINICK, M. M., COHEN, J. D. et CARTER, C. S. (2004). Conflict monitoring and anterior cingulate cortex : an update. *Trends in Cognitive Sciences*, 8(12):539–546.
- BOTVINICK, M. M., NYSTROM, L., FISSELL, K., CARTER, C. S. et COHEN, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, 402(6758):179–181.
- BROWN, J. W. et BRAVER, T. S. (2005). Learned predictions of error likelihood in the Anterior Cingulate Cortex. *Science*, 307:1118–1121.
- BURLE, B., ALLAIN, S., VIDAL, F. et HASBROUCQ, T. (2005). Sequential compatibility effects and cognitive control : Does conflict really matter ? *Journal of Experimental Psychology : Human Perception and Performance*, 31:831–837.
- BURLE, B., BONNET, M., VIDAL, F., POSSAMAÏ, C. A. et HASBROUCQ, T. (2002a). A transcranial magnetic stimulation study of information processing in the motor cortex : relationship between the silent period and the reaction time delay. *Psychophysiology*, 39:207–217.
- BURLE, B., POSSAMAÏ, C. A., VIDAL, F., BONNET, M. et HASBROUCQ, T. (2002b). Executive control in the Simon effect : an electromyographic and distributional analysis. *Psychological Research*, 66:324–336.
- BURLE, B., ROGER, C., ALLAIN, S., VIDAL, F. et HASBROUCQ, H. (2008a). Error negativity does not reflect conflict : A reappraisal of conflict monitoring and anterior cingulate cortex activity. *Journal of Cognitive Neuroscience*, 20(9):1637–55.
- BURLE, B., ROGER, C., VIDAL, F. et HASBROUCQ, T. (2008b). Spatio-temporal dynamics of information processing in the Brain : Recent advances, current limitations and future challenges. *International Journal of Bioelectromagnetism*, 10(1):17–21.
- BURLE, B., VIDAL, F. et BONNET, M. (2004a). Electroencephalographic Nogo potentials in a no-movement context : The case of motor imagery in humans. *Neuroscience Letters*, 360(1-2):77–80.
- BURLE, B., VIDAL, F., TANDONNET, C. et HASBROUCQ, T. (2004b). Physiological evidences for response inhibition in choice reaction time task. *Brain & Cognition*, 56:141–152.
- CALLAWAY, E., HALLIDAY, R., NAYLOR, H. et THOUVENIN, D. (1984). The latency of the average is not the average of the latencies. *Psychophysiology*, 21:571.

- CARBONNELL, L. et FALKENSTEIN, M. (2006). Does the error negativity reflect the degree of response conflict? *Brain Research*, 1095(1):124–130.
- CARBONNELL, L., HASBROUCQ, T., GRAPPERON, J. et VIDAL, F. (2004). Response selection and motor areas : a behavioural and electrophysiological study. *Clinical Neurophysiology*, 115:2164–2174.
- CARTER, C., BOTVINICK, M. M. et COHEN, J. D. (1999). The contribution of the anterior cingulate cortex to executive processes in cognition. *Reviews in the Neurosciences*, 10:49–57.
- CARTER, C. S., BRAVER, T. S., BARCH, D. M., BOTVINICK, M. M., NOLL, D. et COHEN, J. D. (1998). Anterior cingulate cortex, error detection, and the on-line monitoring of performance. *Science*, 280:747–749.
- CARTER, C. S., MACDONALD, A. M., BOTVINICK, M., ROSS, L. L., STENGER, V. A., NOLL, D. et COHEN, J. D. (2000). Parsing executive processes : strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences USA*, 97(4):1944–1948.
- CARTER, C. S., MINTUN, M. et COHEN, J. D. (1995). Interference and facilitation effects during selective attention : an H<sub>2</sub><sup>15</sup>O PET study of stroop task performance. *NeuroImage*, 2(4):264–272.
- CAVANAGH, J. F., COHEN, M. X. et ALLEN, J. J. B. (2009). Prelude to and resolution of an error : EEG phase synchrony reveals cognitive control dynamics during action monitoring. *Journal of Neuroscience*, 29(1):98–105.
- CHAWLA, D., REES, G. et FRISTON, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience*, 2(7):671–676.
- COHEN, A. et SHOUP, R. (1997). Perceptual dimensional constraints in response selection processes. *Cognitive Psychology*, 32(2):128–181.
- COHEN, J. D., BOTVINICK, M. et CARTER, C. S. (2000). Anterior cingulate and prefrontal cortex : who's in control? *Nature Neuroscience*, 3:516–520.
- COHEN, J. D., SERVAN-SCHREIBER, D. et MCCLELLAND, J. L. (1992). A parallel distributed processing approach to automaticity. *American Journal of Psychology*, 105:239–269.
- COHEN, M. X., HELLER, A. S. et RANGANATH, C. (2005). Functional connectivity with anterior cingulate and orbitofrontal cortices during decision-making. *Cognitive Brain Research*, 23(1):61–70.
- COLES, M. G. H., GRATTON, G., BASHORE, T. R., ERIKSEN, C. W. et DONCHIN, E. (1985). A psychophysiological investigation of the continuous flow of human information processing. *Journal of Experimental Psychology : Human Perception and Performance*, 11:529–552.
- COLES, M. G. H., SCHEFFERS, M. K. et HOLROYD, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biological Psychology*, 56:173–189.
- CORBETTA, M., MIEZIN, F. M., DOBMEYER, S., SHULMAN, G. L. et PETERSEN, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, 248(4962):1556–1559.
- COULL, J. T., VIDAL, F., NAZARIAN, B. et MACAR, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science*, 303(5663):1506–1508.
- COURTIÈRE, A., HARDOUIN, J., BURLE, B., VIDAL, F. et HASBROUCQ, T. (2007). Simon effect in the rat : a new model for studying the neural bases of the dual-route architecture. *Behavioural Brain Research*, 179(1):69–75.
- CRAFT, J. L. et SIMON, J. R. (1970). Processing symbolic information from a visual display : interference from irrelevant directional cue. *Journal of Experimental Psychology*, 83:415–420.
- DAVIES, P. L., SEGALOWITZ, S. J. et GAVIN, W. J. (2004a). Development of error-monitoring event-related potentials in adolescents. *Annals of the New York Academy of Sciences*, 1021:324–328.

- DAVIES, P. L., SEGALOWITZ, S. J. et GAVIN, W. J. (2004b). Development of response-monitoring ERPs in 7- to 25-year-olds. *Developmental Neuropsychology*, 25(3):355–376.
- DEBENER, S., ULLSPERGER, M., SIEGEL, M., FIEHLER, K., VON CRAMON, Y. et ENGEL, A. K. (2005). Trial-by-trial coupling of concurrent EEG and fMRI identifies the dynamics of performance monitoring. *Journal of Neuroscience*, 25:11730–11737.
- DEHAENE, S., POSNER, M. et TUCKER, D. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, 5:303–305.
- DELONG, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in Neurosciences*, 13(7):281–285.
- DELORME, A. et MAKEIG, S. (2004). EEGLAB : an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21.
- DI PELLEGRINO, G., CIARAMELLI, E. et LÀDAVAS, E. (2007). The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. *Journal of Cognitive Neuroscience*, 19(2):275–286.
- EGNER, T. et HIRSCH, J. (2005a). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, 8(12):1784–1790.
- EGNER, T. et HIRSCH, J. (2005b). The neural correlates and functional integration of cognitive control in a Stroop task. *NeuroImage*, 24:539–547.
- EICHELE, T., DEBENER, S., CALHOUN, V. D., SPECHT, K., ENGEL, A. K., HUGDAHL, K., von CRAMON, D. Y. et ULLSPERGER, M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proceedings of the National Academy of Sciences USA*, 105(16):6173–6178.
- EMERIC, E. E., BROWN, J. W., LESLIE, M., POUGET, P., STUPHORN, V. et SCHALL, J. D. (2008). Performance monitoring local field potentials in the medial frontal cortex of primates : anterior cingulate cortex. *Journal of Neurophysiology*, 99(2):759–772.
- ENDRASS, T., FRANKE, C. et KATHMANN, N. (2005). Error awareness in a Saccade Countermanding Task. *Journal of Psychophysiology*, 19(4):275–280.
- ENDRASS, T., KLAWOHN, J., SCHUSTER, F. et KATHMANN, N. (2008). Overactive performance monitoring in obsessive-compulsive disorder : ERP evidence from correct and erroneous reactions. *Neuropsychologia*, 46(7):1877–1887.
- ENDRASS, T., REUTER, B. et KATHMANN, N. (2007). ERP correlates of conscious error recognition : aware and unaware errors in an antisaccade task. *European Journal of Neuroscience*, 26(6):1714–1720.
- ERIKSEN, B. A. et ERIKSEN, C. W. (1974). Effects of noise letters upon the identification of target letter in a non-search task. *Perception & Psychophysics*, 16:143–149.
- ERIKSEN, C. W., COLES, M. G. H., MORRIS, L. R. et O'HARA, W. P. (1985). An electromyographic examination of response competition. *Bulletin of the Psychonomic Society*, 23:165–168.
- FALKENSTEIN, M., HOHNSBEIN, J., HOORMANN, J. et BLANKE, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, 78:447–455.
- FALKENSTEIN, M., HOORMANN, J. et HOHNSBEIN, J. (2001). Changes of error-related ERPs with age. *Experimental Brain Research*, 138:258–262.
- FALKENSTEIN, M., HOORMANN, J., CHRIST, S. et HOHNSBEIN, J. (2000). ERP components on reaction errors and their functional significance : a tutorial. *Biological Psychology*, 51:87–107.
- FIEHLER, K., ULLSPERGER, M. et VON CRAMON, Y. (2005). Electrophysiological correlates of error correction. *Psychophysiology*, 42:72–82.



- FITZGERALD, K. D., WELSH, R. C., GEHRING, W. J., ABELSON, J. L., HIMLE, J. A., LIBERZON, I. et TAYLOR, S. F. (2005). Error-related hyperactivity of the anterior cingulate cortex in obsessive-compulsive disorder. *Biological Psychiatry*, 57(3):287–294.
- FRANK, M. J. (2005). Dynamic dopamine modulation in the basal ganglia : a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience*, 17(1):51–72.
- FRANK, M. J., SEEGER, L. C. et O'REILLY, R. C. (2004). By carrot or by stick : cognitive reinforcement learning in parkinsonism. *Science*, 306(5703):1940–1943.
- FRANK, M. J., WOROCH, B. S. et CURRAN, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, 47(4):495–501.
- FRISTON, K. J., HOLMES, A. P., POLINE, J. B., GRASBY, P. J., WILLIAMS, S. C., FRACKOWIAK, R. S. et TURNER, R. (1995). Analysis of fMRI time-series revisited. *NeuroImage*, 2(1):45–53.
- GEHRING, W. J. et FENCSEK, D. (1999). Slamming on the brakes : An electrophysiological study of error response inhibition. Poster presented at the Annual Meeting of the Cognitive Neuroscience Society, April, 11-13.
- GEHRING, W. J. et FENCSEK, D. E. (2001). Functions of the medial frontal cortex in the processing of conflict and errors. *Journal of Neuroscience*, 21:9430–9437.
- GEHRING, W. J., GOSS, B., COLES, M. G. H., MEYER, D. E. et DONCHIN, E. (1993). A neural system for error detection and compensation. *Psychological Science*, 4(6):385–390.
- GEHRING, W. J., HIMLE, J. et NISENSEN, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science*, 11(1):1–6.
- GEHRING, W. J. et KNIGHT, R. T. (2000). Prefrontal-cingulate interactions in action monitoring. *Nature Neuroscience*, 3(5):516–520.
- GEHRING, W. J. et KNIGHT, R. T. (2002). Lateral prefrontal damage affects processing selection but not attention switching. *Cognitive Brain Research*, 13(2):267–279.
- GENTSCH, A., ULLSPERGER, P. et ULLSPERGER, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self- and externally caused failure of goal achievement. *NeuroImage*, 47:2023–2030.
- GRATTON, G., COLES, M. et DONCHIN, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55:468–484.
- GRATTON, G., COLES, M. G. et DONCHIN, E. (1992). Optimizing the use of information : strategic control of activation of responses. *Journal of Experimental Psychology : General*, 121(4):480–506.
- GRATTON, G., COLES, M. G. H., SIREVAAG, E. J., ERIKSEN, C. W. et DONCHIN, E. (1988). Pre- and poststimulus activation of response channels : A psychophysiological analysis. *Journal of Experimental Psychology : Human Perception and Performance*, 14:331–344.
- HABER, S. N. et FUDGE, J. L. (1997). The primate substantia nigra and VTA : integrative circuitry and function. *Critical Reviews in Neurobiology*, 11(4):323–342.
- HADLAND, K. A., RUSHWORTH, M. F. S., GAFFAN, D. et PASSINGHAM, R. E. (2003). The anterior cingulate and reward-guided selection of actions. *Journal of Neurophysiology*, 89(2):1161–1164.
- HAJCAK, G., McDONALD, N. et SIMONS, R. F. (2003). To err is autonomic : error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology*, 40(6):895–903.
- HAJCAK, G., McDONALD, N. et SIMONS, R. F. (2004). Error-related psychophysiology and negative affect. *Brain & Cognition*, 56(2):189–197.

- HAJCAK, G., MOSER, J. S., YEUNG, N. et SIMONS, R. F. (2005a). On the ERN and the significance of errors. *Psychophysiology*, 42(2):151–160.
- HAJCAK, G., NIEUWENHUIS, S., RIDDERINKHOF, K. R. et SIMONS, R. F. (2005b). Error-preceding brain activity : robustness, temporal dynamics, and boundary conditions. *Biological Psychology*, 70(2):67–78.
- HAJCAK, G. et SIMONS, R. F. (2002). Error-related brain activity in obsessive-compulsive undergraduates. *Psychiatry Research*, 110(1):63–72.
- HAJCAK, G. et SIMONS, R. F. (2008). Oops!.. I did it again : an ERP and behavioral study of double-errors. *Brain & Cognition*, 68(1):15–21.
- HASBROUCQ, T., AKAMATSU, M., BURLE, B., BONNET, M. et POSSAMAÏ, C. A. (2000). Changes in spinal excitability during choice reaction time : The H-reflex as a probe of information transmission. *Psychophysiology*, 37:385–393.
- HASBROUCQ, T., POSSAMAÏ, C.-A., BONNET, M. et VIDAL, F. (1999). Effect of the irrelevant location of the response signal on choice reaction time : An electromyographic study in humans. *Psychophysiology*, 36:522–526.
- HAZNEDAR, M., BUCHSBAUM, M., METZGER, M., SOLIMANDO, A., SPIEGEL-COHEN, J. et HOLLANDER, E. (1997). Anterior cingulate gyrus volume and glucose metabolism in autistic disorder. *American Journal of Psychiatry*, 154:1047–1050.
- HEBB, D. O. (1949). *The Organization of Behavior : A Neuropsychological Theory*. Wiley, New-York.
- HEDGE, A. et MARSH, N. (1975). The effect of irrelevant spatial correspondence on two choice response time. *Acta Psychologica*, 39:427–439.
- HEEKEREN, H. R., MARRETT, S., BANDETTINI, P. A. et UNGERLEIDER, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature*, 431(7010):859–862.
- HERRMANN, M. J., RÖMMLER, J., EHLIS, A.-C., HEIDRICH, A. et FALLGATTER, A. J. (2004). Source localization (LORETA) of the error-related-negativity (ERN/Ne) and positivity (Pe). *Cognitive Brain Research*, 20(2):294–299.
- HOGAN, A. M., VARGHA-KHADEM, F., KIRKHAM, F. J. et BALDEWEG, T. (2005). Maturation of action monitoring from adolescence to adulthood : an ERP study. *Developmental Science*, 8(6):525–534.
- HOLROYD, C. et COLES, M. (2002). The neural basis of human error processing : Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109(4):679–709.
- HOLROYD, C., DIEN, J. et COLES, M. (1998). Error-related scalp potentials elicited by hand and foot movements : evidence for an output-independent error-processing system in humans. *Neuroscience Letters*, 242:65–68.
- HOLROYD, C. B., NIEUWENHUIS, S., YEUNG, N., NYSTROM, L., MARS, R. B., COLES, M. G. H. et COHEN, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, 7(5):497–498.
- HOLROYD, C. B., YEUNG, N., COLES, M. G. H. et COHEN, J. D. (2005). A mechanism for error detection in speeded response time tasks. *Journal of Experimental Psychology : General*, 134(2):163–191.
- HOMMEL, B. (1998). Event Files : Evidence for automatic integration of Stimulus-Response episodes. *Visual Cognition*, 5(1/2):183–216.
- HOPFIELD, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA*, 79:2554–2558.
- JIANG, Y. et KANWISHER, N. (2003). Common neural substrates for response selection across modalities and mapping paradigms. *Journal of Cognitive Neuroscience*, 15:1080–1081.

- JUNG, T.-P., MAKEIG, S., WESTERFIELD, M., TOWNSEND, J., COURCHESNE, E. et SEJNOWSKI, T. J. (2001). Analysis and visualizations of single-trial event related potentials. *Human Brain Mapping*, 14:166–185.
- KENNERLEY, S. W., WALTON, M. E., BEHRENS, T. E. J., BUCKLEY, M. J. et RUSHWORTH, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, 9(7):940–947.
- KERNS, J. G. (2006). Anterior cingulate and prefrontal cortex activity in an FMRI study of trial-to-trial adjustments on the Simon task. *NeuroImage*, 33(1):399–405.
- KERNS, J. G., COHEN, J. D., MACDONALD III, A. W., CHO, R. Y., STENGER, V. A. et CARTER, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303:1023–1026.
- KOPP, B., MATTLER, U., GOERTZ, R. et RIST, F. (1996). N2, P3 and the lateralized readiness potential in a nogo task involving selective response priming. *Electroencephalography and Clinical Neurophysiology*, 99:19–27.
- KORNBLUM, S., HASBROUCQ, T. et OSMAN, A. (1990). Dimensional overlap : cognitive basis for stimulus-response compatibility—a model and taxonomy. *Psychological Review*, 97:253–270.
- LADOUCEUR, C. D., DAHL, R. E. et CARTER, C. S. (2007). Development of action monitoring through adolescence into adulthood : ERP and source localization. *Developmental Science*, 10(6):874–891.
- LAMING, D. (1979). Choice reaction performance following an error. *Acta Psychologica*, 43(3):199–224.
- LEVELT, W. J., ROELOFS, A. et MEYER, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1):1–38.
- LIVINGSTONE, M. et HUBEL, D. (1988). Segregation of form, color, movement, and depth : anatomy, physiology, and perception. *Science*, 240(4853):740–749.
- LUU, P., COLLINS, P. et TUCKER, D. M. (2000a). Mood, personality, and self-monitoring : negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology : General*, 129(1):43–60.
- LUU, P., FLAISCH, T. et TUCKER, D. M. (2000b). Medial frontal cortex in action monitoring. *Journal of Neuroscience*, 20(1):464–469.
- MACDONALD, A. W., COHEN, J. D., STENGER, V. A. et CARTER, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288(5472):1835–1838.
- MACLEOD, C. M. (1991). Half a century of research on the Stroop effect : an integrative review. *Psychological Bulletin*, 109(2):163–203.
- MASAKI, H., FALKENSTEIN, M., STÜRMER, B., PINKPANK, T. et SOMMER, W. (2007). Does the error negativity reflect response conflict strength ? Evidence from a Simon task. *Psychophysiology*, 44(4):579–585.
- MASAKI, H. et SEGALOWITZ, S. (2004). Error negativity : A test of the response conflict versus error detection hypotheses. In ULLSPERGER, M. et FALKENSTEIN, M., éditeurs : *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring.*, pages 76–83, Leipzig. MPI of Cognitive Neuroscience.
- MASAKI, H., TANAKA, H., TAKASAWA, N. et YAMAZAKI, K. (2001). Error-related brain potentials elicited by vocal errors. *Neuroreport*, 12(9):1851–1855.
- MATHALON, D. H., BENNETT, A., ASKARI, N., GRAY, E. M., ROSENBLOOM, M. J. et FORD, J. M. (2003). Response-monitoring dysfunction in aging and Alzheimer’s disease : An event-related potential study. *Neurobiology of Aging*, 24(5):675–685.
- MATHALON, D. H., FEDOR, M., FAUSTMAN, W. O., GRAY, M., ASKARI, N. et FORD, J. M. (2002). Response-monitoring dysfunction in schizophrenia : An event-related brain potential study. *Journal of Abnormal Psychology*, 111(1):22–41.

- MAYR, U., AWH, E. et LAUREY, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience*, 6(5):450–452.
- MCCULLOCH, W. S. et PITTS, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133.
- MEYER, D. E., OSMAN, A., IRWIN, D. E. et YANTIS, S. (1988). Modern mental chronometry. *Biological Psychology*, 26:3–67.
- MEYNIER, C., BURLE, B., POSSAMAÏ, C.-A., VIDAL, F. et HASBROUCQ, T. (2009). Neural inhibition and interhemispheric connections in two-choice reaction time : A Laplacian ERP study. *Psychophysiology*, 46:726–730.
- MILTNER, W., BRAUER, J., HECHT, H., TRIPPE, R. et COLES, M. (2004). Parallel brain activity for self-generated and observed errors. In ULLSPERGER, M. et FALKENSTEIN, M., éditeurs : *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring.*, pages 124–129, Leipzig. MPI of Cognitive Neuroscience.
- MILTNER, W. H. R., BRAUN, C. H. et COLES, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task : Evidence for a 'generic' neural system for error-detection. *Journal of Cognitive Neuroscience*, 9:788–798.
- MILTNER, W. H. R., LEMKE, U., WEISS, T., HOLROYD, C., SCHEFFERS, M. K. et COLES, M. G. H. (2003). Implementation of error-processing in the human anterior cingulate cortex : a source analysis of the magnetic equivalent of the error-related negativity. *Biological Psychology*, 64(1-2):157–166.
- MIRENOWICZ, J. et SCHULTZ, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of Neurophysiology*, 72(2):1024–1027.
- MIRENOWICZ, J. et SCHULTZ, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature*, 379(6564):449–451.
- MURPHY, T. I., RICHARD, M., MASAKI, H. et SEGALOWITZ, S. J. (2006). The effect of sleepiness on performance monitoring : I know what I am doing, but do I care ? *Journal of Sleep Research*, 15(1):15–21.
- NAKAMURA, K., ROESCH, M. R. et OLSON, C. R. (2005). Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. *Journal of Neurophysiology*, 93(2):884–908.
- NIEUWENHUIS, S., HOLROYD, C. B., MOL, N. et COLES, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex : origins and functional significance. *Neuroscience and Biobehavioral Reviews*, 28(4):441–448.
- NIEUWENHUIS, S., RIDDERINKHOF, K. R., BLOM, J., BAND, G. P. H. et KOK, A. (2001). Error-related brain potentials are differentially related to awareness of response errors : Evidence from an antisaccade task. *Psychophysiology*, 38:752–760.
- NIEUWENHUIS, S., RIDDERINKHOF, K. R., TALSMA, D., COLES, M. G. H., HOLROYD, C. B., KOK, A. et van der MOLEN, M. W. (2002). A computational account of altered error processing in older age : dopamine and the error-related negativity. *Cognitive, Affective, & Behavioral Neuroscience*, 2(1):19–36.
- NIEUWENHUIS, S., SCHWEIZER, T. S., MARS, R. B., BOTVINICK, M. M. et HAJCAK, G. (2007). Error-likelihood prediction in the medial frontal cortex : a critical evaluation. *Cerebral Cortex*, 17(7):1570–1581.
- NIEUWENHUIS, S., STINS, J. F., POSTHUMA, D., POLDERMAN, T. J. C., BOOMSMA, D. I. et de GEUS, E. J. (2006). Accounting for sequential trial effects in the flanker task : conflict adaptation or associative priming ? *Memory & Cognition*, 34(6):1260–1272.
- NIEUWENHUIS, S. et YEUNG, N. (2005). Neural mechanisms of attention and control : losing our inhibitions ? *Nature Neuroscience*, 8(12):1631–1633.

- O'CONNELL, R. G., DOCKREE, P. M., BELLGROVE, M. A., KELLY, S. P., HESTER, R., GARAVAN, H., ROBERTSON, I. H. et FOXE, J. J. (2007). The role of cingulate cortex in the detection of errors with and without awareness : a high-density electrical mapping study. *European Journal of Neuroscience*, 25(8):2571–2579.
- O'CRAVEN, K. M., ROSEN, B. R., KWONG, K. K., TREISMAN, A. et SAVOY, R. L. (1997). Voluntary attention modulates fMRI activity in human MT-MST. *Neuron*, 18(4):591–598.
- OGAWA, S., LEE, T. M., NAYAK, A. S. et GLYNN, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine*, 14(1):68–78.
- ONTON, J., WESTERFIELD, M., TOWNSEND, J. et MAKEIG, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neuroscience and Biobehavioral Reviews*, 30(6):808–822.
- OVERBEEK, T. J. M., NIEUWENHUIS, S. et RIDDERINKHOF, K. R. (2005). Dissociable components of error processing : On the functional significance of the Pe vis-à-vis the ERN/Ne. *Journal of Psychophysiology*, 19(4):319–329.
- PARDO, J. V., PARDO, P. J., JANER, K. W. et RAICHLE, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences USA*, 87(1):256–259.
- PASCUAL-MARQUI, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sloreta) : technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, 24 Suppl D:5–12.
- PERRIN, F., BERTRAND, O. et PERNIER, J. (1987). Scalp current density mapping : Value and estimation from potential data. *IEEE Transactions on Bio-Medical Engineering*, 34:283–288.
- PERRIN, F., PERNIER, J., BERTRAND, O. et ECHALLIER, J. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, 72:184–187.
- POCHON, J.-B., RIIS, J., SANFEY, A. G., NYSTROM, L. E. et COHEN, J. D. (2008). Functional imaging of decision conflict. *Journal of Neuroscience*, 28(13):3468–3473.
- PRAAMSTRA, P. et SEISS, E. (2005). The neurophysiology of response competition : Motor cortex activation and inhibition following subliminal response priming. *Journal of Cognitive Neuroscience*, 17(3):483–493.
- PROCYK, E., TANAKA, Y. L. et JOSEPH, J. P. (2000). Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nature Neuroscience*, 3(5):502–508.
- QUILODRAN, R., ROTHÉ, M. et PROCYK, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, 57(2):314–325.
- RABBITT, P. M. A. (1966). Errors and error correction in choice reaction tasks. *Journal of Experimental Psychology*, 71:264–272.
- RATCLIFF, R., ZANDT, T. V. et MCKOON, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, 106(2):261–300.
- RIDDERINKHOF, K. R., de VLUGT, Y., BRAMLAGE, A., SPAAN, M., ELTON, M., SNEL, J. et BAND, G. P. H. (2002). Alcohol consumption impairs detection of performance errors in mediofrontal cortex. *Science*, 298:2209–2211.
- RIDDERINKHOF, K. R., NIEUWENHUIS, S. et BASHORE, T. (2003). Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. *Neuroscience Letters*, 348:1–4.
- RIDDERINKHOF, K. R., ULLSPERGER, M., CRONE, E. A. et NIEUWENHUIS, S. (2004a). The role of the medial frontal cortex in cognitive control. *Science*, 306:443–447.
- RIDDERINKHOF, K. R., van den WILDENBERG, W. P. M., SEGALOWITZ, S. J. et CARTER, C. S. (2004b). Neurocognitive mechanisms of cognitive control : the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain & Cognition*, 56(2):129–140.

- RODRÍGUEZ-FORNELLS, A., KURZBUCH, A. R. et MÜNTE, T. F. (2002). Time course of error detection and correction in humans : Neurophysiological evidence. *Journal of Neuroscience*, 22:9990–9996.
- ROELOFS, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42(1-3):107–142.
- ROELOFS, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64(3):249–284.
- ROELOFS, A. (2003). Goal-referenced selection of verbal action : modeling attentional control in the Stroop task. *Psychological Review*, 110(1):88–125.
- ROELOFS, A. et HAGOORT, P. (2002). Control of language use : cognitive modeling of the hemodynamics of Stroop task performance. *Cognitive Brain Research*, 15(1):85–97.
- ROELOFS, A., van TURENNOUT, M. et COLES, M. G. H. (2006). Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. *Proceedings of the National Academy of Sciences USA*, 103(37):13884–13889.
- ROGER, C., HASBROUCQ, T., RABAT, A., VIDAL, F. et BURLE, B. (2009). Neurophysics of temporal discrimination in the rat : A mismatch negativity study. *Psychophysiology*, 46:1028–1032.
- ROMO, R. et SCHULTZ, W. (1990). Dopamine neurons of the monkey midbrain : contingencies of responses to active touch during self-initiated arm movements. *Journal of Neurophysiology*, 63(3):592–606.
- ROSENBLATT, F. (1958). The perceptron : a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408.
- RÖSLER, F. et FINGER, T. (1993). A psychophysiological analysis of response-channel activation and outcome states in Eriksen’s noise-compatibility task. *Psychological Research*, 55:20–28.
- RUCHSOW, M., GROTHE, J., SPITZER, M. et KIEFER, M. (2002). Human anterior cingulate cortex is activated by negative feedback : evidence from event-related potentials in a guessing task. *Neuroscience Letters*, 325(3):203–206.
- RUMELHART, D. E., HINTON, G. E. et WILLIAMS, R. J. (1986a). Learning internal representation by error propagation. In RUMELHART, D. E. et MCCLELLAND, J. L., éditeurs : *Parallel Distributed Processing. Exploration of the Microstructure of Cognition.*, volume 1, pages 318–362. MIT Press, Cambridge.
- RUMELHART, D. E., HINTON, G. E. et WILLIAMS, R. J. (1986b). Learning representations by back-propagating errors. *Nature*, 323:533–536.
- RUSHWORTH, M. F. S., BEHRENS, T. E. J., RUDEBECK, P. H. et WALTON, M. E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends in Cognitive Sciences*, 11(4):168–176.
- SANTESSO, D. L., SEGALOWITZ, S. J. et SCHMIDT, L. A. (2006). Error-related electrocortical responses in 10-year-old children and young adults. *Developmental Science*, 9(5):473–481.
- SATOH, T., NAKAI, S., SATO, T. et KIMURA, M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. *Journal of Neuroscience*, 23(30):9913–9923.
- SCHEFFERS, M. K. et COLES, M. G. (2000). Performance monitoring in a confusing world : error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology : Human Perception and Performance*, 26(1):141–151.
- SCHEFFERS, M. K., COLES, M. G. H., BERNSTEIN, P., GEHRING, W. J. et DONCHIN, E. (1996). Event-related brain potentials and error-related processing : An analysis of incorrect responses to go and no-go stimuli. *Psychophysiology*, 33:42–53.
- SCHULTZ, W. (1986). Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *Journal of Neurophysiology*, 56(5):1439–1461.

- SCHULTZ, W. (2002). Getting formal with dopamine and reward. *Neuron*, 36(2):241–263.
- SCHULTZ, W., APICELLA, P. et LJUNGBERG, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, 13(3):900–913.
- SCHULTZ, W., DAYAN, P. et MONTAGUE, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.
- SCHUMACHER, E. H. et D’ESPOSITO, M. (2002). Neural implementation of response selection in humans as revealed by localized effects of stimulus-response compatibility. *Human Brain Mapping*, 17:193–201.
- SEGALOWITZ, S. J. et DAVIES, P. L. (2004). Charting the maturation of the frontal lobe : an electrophysiological strategy. *Brain & Cognition*, 55(1):116–133.
- SMID, H. G. O. M., MULDER, G. et MULDER, L. J. M. (1990). Selective response activation can begin before stimulus recognition is complete : A psychophysiological and error analysis of the continuous flow. *Acta Psychologica*, 74:169–201.
- SMITH, P. L. et RATCLIFF, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27(3):161–168.
- STAUDE, G., FLACHENECKER, C., DAUMER, M. et WOLF, W. (2001). Onset detection in surface electromyographic signals : A systematic comparison of methods. *Journal on Applied Signal Processing*, 2:67–81.
- STEMMER, B., VIHLA, M. et SALMELIN, R. (2004). Activation of the human sensorimotor cortex during error-related processing : A magnetoencephalography study. *Neuroscience Letters*, 362:44–47.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18:643–662.
- STÜRMER, B., LEUTHOLD, H., SOETENS, E., SCHROTER, H. et SOMMER, W. (2002). Control over location-based response activation in the Simon task : behavioral and electrophysiological evidence. *Journal of Experimental Psychology : Human Perception and Performance*, 28(6):1345–1363.
- SUTTON, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, 3:9–44.
- SUTTON, R. S. et BARTO, A. G. (1990). Time-derivative models of pavlovian reinforcement. In GABRIEL, M. et MOORE, J., éditeurs : *Learning and computational neuroscience : Foundations of adaptive networks*, pages 497–537. Cambridge, MIT Press.
- SUTTON, R. S. et BARTO, A. G. (1998). Reinforcement Learning : An introduction. *IEEE Transactions on Neural Networks / a publication of the IEEE Neural Networks Council*, 9(5):1054–1054.
- SUTTON, R. S., BARTO, A. G. et WILLIAMS, R. J. (1992). Reinforcement learning is direct adaptive optimal control. *IEEE control systems magazine (New York, N.Y. : 1985)*, 12(2):19–22.
- SWICK, D., et TURKEN, U. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proceedings of the National Academy of Sciences USA*, 99(25):16354–16359.
- TANDONNET, C., BURLE, B., HASBROUCQ, T. et VIDAL, F. (2005). Spatial enhancement of EEG traces by surface Laplacian estimation : comparison between local and global methods. *Clinical Neurophysiology*, 116(1):18–24.
- TANDONNET, C., BURLE, B., VIDAL, F. et HASBROUCQ, T. (2003). The influence of time preparation on motor processes assessed by surface Laplacian estimation. *Clinical Neurophysiology*, 114:2376–2384.
- TANIGUCHI, Y., BURLE, B., VIDAL, F. et BONNET, M. (2001). Deficit in motor cortical activity for simultaneous bimanual responses. *Experimental Brain Research*, 137:259–268.

- TOOTELL, R. B., REPPAS, J. B., KWONG, K. K., MALACH, R., BORN, R. T., BRADY, T. J., ROSEN, B. R. et BELLIVEAU, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, 15(4):3215–3230.
- ULLSPERGER, M., BYLSMA, L. M. et BOTVINICK, M. M. (2005). The conflict adaptation effect : it's not just priming. *Cognitive, Affective, & Behavioral Neuroscience*, 5(4):467–472.
- ULLSPERGER, M. et VON CRAMON, D. Y. (2001). Subprocesses of performance monitoring : a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage*, 14:1387–1401.
- URSU, S., STENGER, V. A., SHEAR, M. K., JONES, M. R. et CARTER, C. S. (2003). Overactive action monitoring in obsessive-compulsive disorder : evidence from functional magnetic resonance imaging. *Psychological Science*, 14(4):347–353.
- USHER, M. et MCCLELLAND, J. L. (2001). The time course of perceptual choice : The leaky, competing accumulator model. *Psychological Review*, 108(3):550–592.
- VALLE-INCLÁN, F. et REDONDO, M. (1998). On the automaticity of ipsilateral response activation in the Simon effect. *Psychophysiology*, 35:366–371.
- VAN BOXTEL, G. J. M., GERAATS, L. H. D., VAN DEN BERG-LESSEN, M. M. C. et BRUNIA, C. H. M. (1993). Detection of EMG onset in ERP research. *Psychophysiology*, 30:405–412.
- VAN SCHIE, H. T., MARS, R. B., COLES, M. G. H. et BEKKERING, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, 7(5):549–554.
- VAN 'T ENT, D. et APKARIAN, P. (1999). Motoric response inhibition in finger movement and saccadic eye movement : a comparative study. *Clinical Neurophysiology*, 110(6):1058–1072.
- VAN VEEN, V. et CARTER, C. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, 14(4):593–602.
- VIDAL, F., BURLE, B., BONNET, M., GRAPPERON, J. et HASBROUCQ, T. (2003a). Error negativity on correct trials : a reexamination of available data. *Biological Psychology*, 64:265–282.
- VIDAL, F., GRAPPERON, J., BONNET, M. et HASBROUCQ, T. (2003b). The nature of unilateral motor commands in between-hands choice tasks as revealed by surface Laplacian estimation. *Psychophysiology*, 40:796–805.
- VIDAL, F., HASBROUCQ, T., GRAPPERON, J. et BONNET, M. (2000). Is the “error negativity” specific to errors ? *Biological Psychology*, 51:109–128.
- VLAMINGS, P. H. J. M., JONKMAN, L. M., HOEKSMAN, M. R., van ENGELAND, H. et KEMNER, C. (2008). Reduced error monitoring in children with autism spectrum disorder : an ERP study. *European Journal of Neuroscience*, 28(2):399–406.
- VOCAT, R., POURTOIS, G. et VUILLEUMIER, P. (2008). Unavoidable errors : A spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia*, 46(10):2545–2555.
- WATSON, J. D., MYERS, R., FRACKOWIAK, R. S., HAJNAL, J. V., WOODS, R. P., MAZZIOTTA, J. C., SHIPP, S. et ZEKI, S. (1993). Area V5 of the human brain : evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, 3(2):79–94.
- WEST, R. (2004). The effects of aging on controlled attention and conflict processing in the Stroop task. *Journal of Cognitive Neuroscience*, 16(1):103–113.
- WILLEMSSEN, R., MÜLLER, T., SCHWARZ, M., FALKENSTEIN, M. et BESTE, C. (2009). Response monitoring in de novo patients with Parkinson's disease. *PLoS ONE*, 4(3):e4898.



- WILLIAMS, S. M. et GOLDMAN-RAKIC, P. S. (1993). Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cerebral Cortex*, 3(3):199–222.
- WINER, B. J. (1971). *Statistical principles in experimental design : Design and analysis of factorial experiments*. McGraw-Hill, NewYork.
- YEUNG, N. (2004). Relating cognitive and affective theories of the error-related negativity. In ULLSPERGER, M. et FALKENSTEIN, M., éditeurs : *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring.*, pages 63–70, Leipzig. MPI of Cognitive Neuroscience.
- YEUNG, N., BOTVINICK, M. M. et COHEN, J. D. (2004). The neural basis of error detection : Conflict monitoring and the error-related negativity. *Psychological Review*, 111(4):931–959.
- YORDANOVA, J., FALKENSTEIN, M., HOHNSBEIN, J. et KOLEV, V. (2004). Parallel systems of error processing in the brain. *NeuroImage*, 22:590–602.
- ZEKI, S., WATSON, J. D., LUECK, C. J., FRISTON, K. J., KENNARD, C. et FRACKOWIAK, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience*, 11(3):641–649.



# Table des figures

1.1	Exemple d'une tâche de TR de choix à deux éventualités . . . . .	12
1.2	Illustration d'un ajustement post-erreur . . . . .	13
1.3	Tâche de Simon manipulant la compatibilité spatiale . . . . .	14
1.4	Illustration des l'effets séquentiels sur le TR . . . . .	15
1.5	Exemple d'enregistrement de l'activité EMG des muscles impliqués dans un essai correct avec ébauche d'erreur . . . . .	16
1.6	Coupe sagittale médiane du cerveau . . . . .	17
1.7	Variations de la fréquence de décharge des neurones DA au moment de l'attribution de la récompense . . . . .	22
1.8	Dynamique temporelle, topographie et localisation de onde d'erreur ( $N_E$ /ERN) . .	23
1.9	Dynamique temporelle de la négativité liée au feedback (FRN) . . . . .	24
1.10	Dynamique temporelle et topographie de la "négativité de l'observateur" . . . . .	25
1.11	Dynamique temporelle de la Positivité d'Erreur ( $P_E$ ) . . . . .	27
1.12	Dynamique temporelle de l'activité EEG après calcul du Laplacien de surface au dessus des cortex moteurs et au dessus de l'AMS/CCA . . . . .	28
1.13	Dynamique temporelle de la $N2$ calée sur le stimulus . . . . .	29
1.14	Dynamique temporelle de la $N2$ calée sur la réponse . . . . .	30
1.15	Grand moyennage des tracés EEG monopolaires pour les erreurs et les essais corrects en fonction de la consigne (vitesse ou précision) . . . . .	31
1.16	Grands moyennages des tracés EEG après calcul du Laplacien de surface dans les erreurs, les essais corrects et les essais corrects avec ébauche d'erreur . . . . .	33
1.17	Tableau résumé des marqueurs du contrôle cognitif dans les différentes pathologies	42
2.1	Schéma simplifié d'un neurone formel au sein d'un réseau de neurones . . . . .	44

2.2	Schéma détaillé d'un neurone formel . . . . .	44
2.3	Exemples de réseaux de neurones. . . . .	45
2.4	Schéma représentant l'hypothèse de la Mismatch . . . . .	47
2.5	Réalisation de la tâche d'Eriksen (Eriksen et Eriksen, 1974) . . . . .	49
2.6	Modèle du Conflit de Botvinick <i>et al.</i> (2001) adapté pour la tâche d'Eriksen . . . .	51
2.7	Simulation d'un essai correct et d'une erreur. . . . .	55
2.8	Les activités $N_E$ et $N_2$ observées et comparées aux simulations du modèle du conflit par Yeung <i>et al.</i> (2004) . . . . .	56
2.9	Exemples de stimuli utilisés par Egner et Hirsch (2005a) dans leur étude IRMf . .	57
2.10	Activités EMG moyennées dans les erreurs et les essais corrects (Allain <i>et al.</i> , 2004b)	59
2.11	Schéma du modèle WEAVER ++ adapté à la tâche de Stroop . . . . .	61
2.12	Schéma du modèle WEAVER ++ avec le CCA . . . . .	62
2.13	Simulation BOLD, modèle de Roelofs (Roelofs et Hagoort, 2002) . . . . .	64
2.14	Illustration de la tâche et du modèle utilisés par Roelofs <i>et al.</i> (2006) . . . . .	66
2.15	Modèle ERN-RL simplifié proposé par Holroyd <i>et al.</i> (2005) . . . . .	68
2.16	Modèle ERN-RL proposé par Holroyd <i>et al.</i> (2005) adapté pour la tâche d'Eriksen	70
2.17	Modélisation du fonctionnement des GDB utilisée par Frank (2005) pour la construc- tion du modèle neuromimétique des GDB. . . . .	73
2.18	Tableau représentant les différents aspects des modèles neuromimétiques . . . . .	78
3.1	Examples of "partial errors" obtained in the experiment and in the simulation . .	84
3.2	Comparison between grand-averaged $N_E$ and conflict . . . . .	90
3.3	ERP-Images of the $N_E$ and of the simulated conflict . . . . .	92
3.4	Classes analysis of the amplitude of the $N_E$ and the amount of conflict . . . . .	93
3.5	Amount of incorrect response activation as a function of $\Delta$ . . . . .	94
3.6	Empirical co-activation and conflict measured on the averaged traces . . . . .	96
3.7	Single trial analysis of response activation . . . . .	97
3.8	Topographies and time-courses of the activities under interest . . . . .	98
3.9	Impact of interpolation on the time-course of activity of interest . . . . .	109
4.1	Grand average (FCz) of the three categories of trials for the monopolar and the Laplacian transformed data . . . . .	119
4.2	Example of decomposition and component selection for subject 01 . . . . .	120
4.3	Topographies of the $Ne_{IC_S}$ for each of the 10 subjects . . . . .	122

4.4	Grand average of the $Ne_{IC_S}$ time course and of the Laplacian transformed data .	122
4.5	Source localization of the $Ne_{IC_S}$ . . . . .	123
4.6	Source localization results obtained with sLORETA . . . . .	124
4.7	Grand average of the fronto-lateral negativity observed over FC1/FC2 after Laplacian computation for errors, partial errors and correct trials . . . . .	126
4.8	Topographies and time courses of the medial and lateral activity for the three types of trials . . . . .	126
5.1	Illustration de la tâche de Simon dans sa version dite “accessoire” . . . . .	134
5.2	Placement des électrodes EMG en regard du <i>flexor policis brevis</i> . . . . .	135
5.3	Expérience 1 : Résultats Comportementaux . . . . .	136
5.4	Expérience 1 : Dynamique temporelle et topographie de la N-40 . . . . .	138
5.5	Expérience 1 : Dynamique temporelle et topographie du patron d’activation-inhibition	140
5.6	Expérience 1 : Dynamique temporelle de l’activité EEG au dessus des cortex pré-moteurs . . . . .	141
5.7	Expérience 1 : Résumé des résultats . . . . .	144
5.8	Expérience 2 : Résultats Comportementaux . . . . .	145
5.9	Expérience 2 : Dynamique temporelle de la N-40 et du patron d’activation-inhibition	147
5.10	Expérience 2 : Dynamique temporelle de l’activité EEG au dessus des cortex pré-moteurs . . . . .	149
5.11	Expérience 2 : Résumé des résultats . . . . .	151
6.1	Expérience IRMf - Design expérimental . . . . .	162
6.2	Expérience IRMf - Résultats comportementaux . . . . .	165
6.3	Expérience IRMf - Table des activations dans le contraste Incompatible-Compatible	166
6.4	Expérience IRMf - Résultats métaboliques dans le contraste Incompatible-Compatible	167
6.5	Expérience IRMf - Table des activations dans le contraste Compatible-Incompatible	167
6.6	Expérience IRMf - Résultats métaboliques - Activités temporo-occipitales . . . . .	168
7.1	Simulation dipolaire de la $N2$ calée sur le stimulus et de la $P3$ calée sur la réponse	180



## Annexe A

# Papiers non directement liés à la thèse

## Neurophysics of temporal discrimination in the rat: A mismatch negativity study

CLÉMENCE ROGER,<sup>a</sup> THIERRY HASBROUCQ,<sup>a</sup> ARNAUD RABAT,<sup>b</sup> FRANCK VIDAL,<sup>a</sup> AND  
BORIS BURLE<sup>a</sup>

<sup>a</sup>Laboratoire de Neurobiologie de la Cognition, Aix-Marseille Université, CNRS, Marseille, France

<sup>b</sup>Institut de Médecine Navale du Service de Santé des Armées, Toulon, France

### Abstract

Behavioral estimates of time discrimination threshold on animals might be contaminated by the conditioning procedure used and by attentional effects. To avoid such side effects, we measured time discrimination by recording the rat electroencephalographic response to small temporal variations. Freely moving rats were presented with repetitive sounds, some of them being occasionally shorter than the standard, to produce a Mismatch Negativity (MMN) which is known to primarily involve preattentive processes. The smallest difference eliciting a MMN located the discrimination threshold between 16% and 33% of the standard, without attentional confound. Being observed in several species, MMN can be used to decipher both the phylogenetic and ontogenetic evolution of time discrimination, without attentional confound.

**Descriptors:** Time estimation, Mismatch negativity, Rats

Time estimation is critical to react adaptively to changing environments. Understanding the architecture of the “time processor” is a major challenge in cognitive neuroscience (see Buhusi & Meck, 2005, for a recent overview). Although timing research in humans has largely expanded in the last few years (see, e.g., Burle & Casini, 2001; Coull, Vidal, Nazarian, & Macar, 2004; Wearden, 2003), most behavioral and physiological knowledge relative to time estimation is based on animal data. As a matter of fact, the major models, “Scalar Expectancy Theory” (Gibbon, 1977) and the “Striatal Beat-Frequency” model (Matell & Meck, 2004) were inspired by rat data, before being extended to humans.

Time discrimination assessment in animals, as opposed to human subjects, normally relies on lengthy conditioning procedures. In addition to the time-consuming aspect, conditioning is not immune to problems. First, in non-human species, temporal performance largely depends on the conditioning procedure employed: For example, Lejeune and Jasselette (1986) showed that time discrimination performances of pigeons changed dramatically depending on the action the pigeons had to perform to express their responses (perching vs. treadle pressing). Second,

timing performance is highly sensitive to the amount of attention devoted to time (Burle & Casini, 2001; Casini & Macar, 1997; Coull et al., 2004; Thomas & Weaver, 1975), a factor very difficult to control on animals, which makes inter-species comparisons difficult. Indeed, although timing performance across species presents a monotonic degradation as we go down the phylogenetic scale (Lejeune & Wearden, 1991), this may not necessarily mean, however, that temporal discriminability *per se* is degraded, since differences might be due, at least partly, to a reduction in attentional capacities as a function of species.

The goal of the present study was to develop a methodology allowing to measure, without such biases, time discrimination in rats, and that could also be used for other species (including humans). To do so, we measured the rat brain response to sounds varying in durations: When subjects are presented with repetitive (standard) sounds, the occasional occurrence of a sound varying along one dimension (deviant) induces a specific brain response, the “Mismatch Negativity” (MMN, see Näätänen, Paavilainen, Rinne, & Alho, 2007, for an overview). This brain response mainly recruits preattentive processes, since its elicitation does not depend on any behavioral training or intention of the animal to discriminate between stimuli (see Sussman, 2007, for a recent review). MMN has been used to estimate pitch discrimination threshold in humans: When the difference in pitch between the deviants and the standards was below the behaviorally estimated threshold, no MMN was elicited (Sams, Paavilainen, Alho, & Näätänen, 1985). This methodology has also been applied to study temporal discrimination in humans (Jacobsen & Schröger, 2003; Jaramillo, Paavilainen, & Näätänen, 2000). It is to be noted that, since no behavioral task is required, this methodology can be applied on populations in which behavioral

The authors wish to thank F. Macar and L. Casini for helpful comments on timing, B. Poucet and A. Norena for sharing their expertise in rat electrophysiology, and R. Pernaud, Dany Paleressompoulle, and Christine Manrique for technical assistance. This research was supported by a doctoral grant from the French Ministry of Research to C.R. and from a research grant from CNRS “Cognition et traitement de l’information,” CTI 02-09.

Send reprint requests to: Dr. Boris Burle, Laboratoire de Neurobiologie de la Cognition, Université de Provence (CNRS), Case C, 3, Place Victor Hugo, 13331 Marseille, Cedex 3, France. E-mail: boris.burle@univ-provence.fr



testing would be complex, like infants (Brannon, Roussel, Meck, & Woldorff, 2004).

Since MMNs to duration have already been observed in mice (Umbricht, Vysotski, Latanov, Nitsch, & Lipp, 2005) and MMNs to pitch have been reported in rats (Eriksson & Villa, 2005; Ruusuvirta, Penttonen, & Korhonen, 1998), it seemed feasible to use this methodology to study time processing in this latter species. We thus searched for the smallest time difference between the standard and the deviant that elicits a MMN, indicating the time discrimination threshold in rats.

## Material and Methods

### Participants

Ten Long Evans rats, weighing 350–400 g served as subjects. All procedures concerning animals were in accordance with the guidelines of the French Ministry of Agriculture and of the National Commission of Animal Experimentation.

### Stimuli

A MMN to duration can be elicited in two ways: either by varying the duration of the sound itself (Jaramillo et al., 2000) or by varying the interval between two sounds (inter-trials interval, ITI) (Brannon et al., 2004; Kujala, Kallio, Tervaniemi, & Näätänen, 2001). In the present study, since we wanted to use large deviance (up to 70%) to maximize the chances of getting an MMN, we opted for the first method, even if it may have led to suboptimal compromises (Jacobsen & Schröger, 2003, see Discussion), since the second one would have required much longer ITI (keeping the smallest ITI at at least 300 ms would require a standard ITI of 1 s), considerably increasing the total duration of the recording sessions.

Standard and deviant stimuli were 3 kHz frequency sounds, with an intensity set at 66 dB, presented by a buzzer positioned at the rat head-level. They differed only in duration. Standard tones lasted 150 ms, whereas deviant tones were shorter. Five levels of deviation were used: 125 ms (deviance: 16.67%, 125-deviant), 100 ms (deviance: 33.33%, 100-deviant), 75 ms (deviance: 50%, 75-deviant), 50 ms (deviance: 66.67%, 50-deviant). (We also tried 25 ms deviants, but the dynamics of the buzzer did not allow such a short duration to produce well formed sounds, hence inducing inconsistent evoked potentials. They were not analyzed further.) Only one deviant was used on each session, and the different deviants were delivered on different experimental sessions. They were presented in pseudo-randomized sequences and occurred with a .2 probability. The interval between the end of the stimulus and the onset of the next one was 500 ms. In each session, 200 standard and 50 deviant stimuli were presented. Each rat ran 6 sessions for each deviant condition. They were freely moving in an empty experimental box (31 cm × 42 cm × 40 cm) adapted for electroencephalographic (EEG) recordings. EEG activities were recorded continuously during presentation of stimulus series.

### Surgery and Electrophysiological Recordings

For the surgery, rats were anesthetized with a solution of ketamine (Ketamine 1000, Virbac, Carros, France, 62.5 mg/Kg) and medetomidine (Domitor, Orion Pharma, Espoo, Finland, 0.4 mg/Kg) injected intramuscularly and positioned in a stereotaxic frame (David Kopf Instrument, Tujunga, CA) with the incisor bar set at 3.3 mm below the interaural line. The recordings were

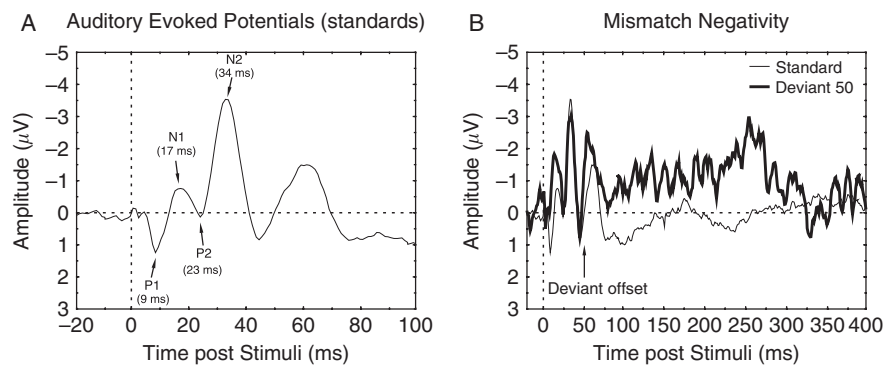
performed at minimum one week after surgery, while the animals were alert and freely moving.

Electrophysiological recordings were performed with an Active 2 system (BioSemi, Amsterdam, The Netherlands) adapted to rat. Electrodes were stainless steel epidural screws (head diameter: 2.5 mm; shaft diameter: 1.57 mm; shaft length: 1.6 mm). We placed 5 electrodes on the rat skull following Paxinos and Watson (1986) coordinates above the two primary motor cortices (bregma AP+1.2 and L ± 2.5), the two parietal cortices (bregma AP−4 and L ± 2.5) and the anterior cingulate cortex (bregma AP+3.5 and L = 0). The reference electrode was located at position bregma AP+4.5 and L+1.6. Two additional electrodes were located at AP+4.5 and L−1.6, and AP+5 and L0, to serve as active references (see BioSemi Web site—<http://www.biosemi.com>—for more precisions). The EEG data were recorded continuously (sampling rate: 1024 Hz, filters: DC to 268 Hz, 3 dB/octave) and saved on the computer disk for offline analysis (performed with BrainAnalyser Brain Products, Munich, Germany). Artifacts were removed by visual inspection of all EEG traces: the rejection mainly concerned high frequency waves (≈100–200 Hz, ripples) occurring primarily during grooming (see Buzsáky, 2006). After artifacts rejection, the correct traces were filtered (band-pass 2–20 Hz), segmented and averaged time-locked to the stimulus onset, for each type of stimuli (standard and the four deviants) separately, to reveal the auditory evoked potentials. Baseline was taken from −20 to 0 ms before stimulus onset (Ruusuvirta et al., 1998).

## Results

For one rat, the recordings at parietal electrodes were very noisy, impeding analysis. Its data were discarded and the analyses were performed on the 9 remaining rats. Figure 1A presents the grand-averaged evoked potentials for the standard recorded at parietal sites. One can easily observe the early and middle latencies Auditory Evoked Potentials (AEP), with the typical P10 (latency = 9 ms), N17 (17 ms), P23 (24 ms), and N38 (34 ms) components (Knight, Brailowsky, Scabini, & Simpson, 1985). For the deviants, a sustained negativity follows those AEPs (Figure 1B).

Since previous studies have reported the presence of a MMN starting about 15–20 ms after stimuli offset (Ruusuvirta et al., 1998), we first compared the mean signal amplitude obtained in the 25–75 ms time windows after the offset of each deviant to the equivalent one for the standard (i.e., from 175 to 225 ms). A repeated measures ANOVA demonstrated a significant main effect of the stimulus duration ( $F(4,32) = 6.94$ , Huynh-Feldt  $\epsilon = 0.85$ ,  $p < .002$ ). Contrast analyses revealed that surface was significantly larger for the deviant than for the standard condition for the 50-deviant ( $F(1,8) = 13.54$ ,  $p < .01$ ), the 75-deviant ( $F(1,8) = 38.1$ ,  $p < .001$ ), and the 100-deviant ( $F(1,8) = 5.11$ ,  $p = .05$ ). The difference was far from significance for the 125-deviant conditions ( $F(1,8) = 0.75$ ,  $p = .41$ ). To better characterize the brain responses, we fitted a logistic function (often used to fit behavioral and neuronal psychophysics data, see, for example, de Lafuente & Romo, 2005) to the amplitudes. As can be seen in Figure 2, the amplitude of the brain response nicely follows the logistic function (residual mean squares = 0.03), showing that the brain responses decrease non-linearly as the sound durations converge toward the standard, expressing a form of Weber's law. We also estimated the difference between the standard and the deviants in the 25–75 ms time interval after offset of the deviants,



**Figure 1.** (A) Grand-average auditory evoked potentials. One can clearly observe the middle-latency component of the rat auditory evoked potentials as defined by Knight et al. (1985), namely the P1–N1 and the P2–N2 complex. (B) Grand-average for the standards (thin line) and for the 50-deviants (thick line). Shortly after the offset of the deviant, one can see a sustained negativity, which likely correspond to the MMN on rat.

and these differences were compared to 0 with student *t*-tests. The results confirmed the one reported above: the difference between deviants and standard was significant for the 50-deviant ( $t(8) = 2.5$ ;  $p < .05$ ), the 75-deviant ( $t(8) = 3.2$ ;  $p < .05$ ) and the 100-deviant ( $t(8) = 3.1$ ;  $p < .05$ ), but not for the 125-deviant ( $t(8) = 1.2$ ;  $p = .28$ ).

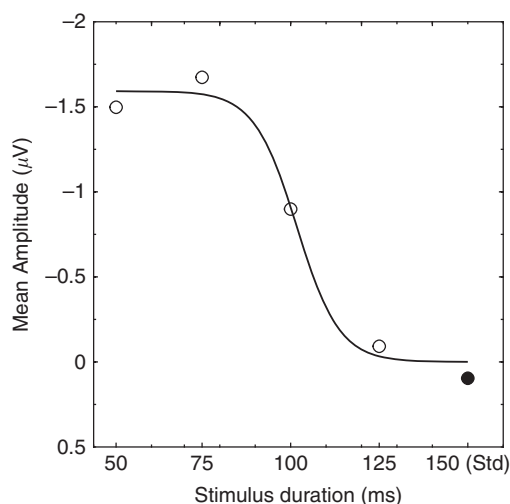
## Discussion

The issue of how animals and humans accurately estimate time has been a long standing challenge in behavioral, and more recently, brain sciences (see Buhusi & Meck, 2005, for a recent review). Time discrimination is normally inferred from the performance obtained in behavioral tasks. However, the actual performance is also affected by non-timing factors, such as the conditioned task and the conditioning procedure (Lejeune & Jasselette, 1986), and by other cognitive factors, like attention

(Burle & Casini, 2001; Coull et al., 2004; Thomas & Weaver, 1975). Those factors likely alter the reliability of our estimation of the actual time discrimination capabilities. Here, by measuring the brain response to temporal variations without resorting to conditioning procedures, we sought to obtain a less biased estimate of time discrimination. We indeed observed a difference between standard and deviants that shows up only when the difference in durations was above 16%. This activity shares all the basic properties of the MMN recorded on humans. First, the MMN-like is elicited despite the fact that the sounds were completely irrelevant and that rats had no reason to pay attention in any way to them (no reinforcement nor punishment). Second, the timing of the difference obtained in the present study corresponds to the one obtained in rats for frequency discrimination (Ruusuvirta et al., 1998), and fits with that obtained in humans. Indeed, in humans the MMN occurs in the N2 latency range, as also observed here for the MMN-like. We will thus consider in the following that it is a MMN.

The results show that although the rat's brain can discriminate 100 ms from 150 ms, it cannot differentiate 125 ms from 150 ms. Accordingly, the discrimination threshold lies between 16% and 33%, which fits behavioral estimates of this threshold around 25% (Lejeune & Wearden, 1991). Note that, although a MMN was clearly present for the 100-deviant, it was of smaller amplitude compared to the 50- and 75-deviant (see, however, Horváth et al., 2008). The fit of a logistic function to the amplitude of the response (Figure 2) confirms that the amplitude of the brain response follows a form of Weber's law as reported on humans (Brannon, Libertus, Meck, & Woldorff, 2008).

The present results were obtained for short durations. There is, however, a debate in the time estimation literature on whether "short" (<200ms) and "long" (>200ms) durations share the same processing operations. It is indeed often considered that estimation of "short" and "long" durations are processed by different networks. For example, Rammsayer, Hennig, Haag, and Lange (2001) reported differential effects of a noradrenergic agonist (reboxine) for second and subsecond judgments. Recently, Lee et al. (2008) reported that repetitive transcranial magnetic stimulation of the cerebellum affects "short" intervals but not "long" ones. This distinction is, however, still debated (see, e.g., Rammsayer & Ulrich, 2005). Although this debate is of fundamental importance, it is somehow orthogonal to our concern. Indeed, our goal was to set up a situation allowing us to estimate time discrimination while getting rid of some potential



**Figure 2.** Amplitude of the EEG response in a time-window between 25 and 100 ms after sound offset, as a function of stimulus duration ( $\circ$  = deviants,  $\bullet$  = standard). The solid lines present the logistic function (often used as a psychophysical function) adjusted to those amplitudes. The amplitudes nicely show an S-shaped evolution as the sounds converge toward the standard.

artifacts. In this respect, it is to be noted that both “short” and “long” intervals are sensitive to attentional processes. Indeed, in their original demonstration of the effect of cognitive load on time perception, Thomas and Weaver (1975) used very short stimuli, of 40 and 80 ms. Recently, Rammsayer and Ulrich (2005) observed similar effects for 100 ms. Thus, at minimum, the present data open the possibility to study the processing of short durations without attentional and conditioning confounds.

As already presented (see “Materials and Methods” section), a MMN to durations deviance can be elicited in two ways (deviance in the durations themselves or in the ITI). For this first attempt to establish a discrimination curve based on MMN, we wished to be on the safe side by using large deviance (about 70%). Varying the ITI down to 70% of the standard would have required rather long standards ITI (at least 1 s, to get a minimal ITI value of 300 ms), making the experiment much longer. To keep the experiment reasonably short (each rat already ran 30 recording sessions), we made the choice of changing stimuli durations, instead of inter-trials durations. We also decided not to reverse the standard and deviant, since this would have required at least 10 new sessions, even if this may have led to non-optimal compromises (Jacobsen & Schröger, 2003). Despite those limitations, the present data already demonstrate the usability of the MMN as a tool to estimate the time discrimination threshold on non-human species.

Extending those results to longer durations would, however, necessitate varying ITI. Indeed, when the durations of the standard extend beyond  $\approx 300$  ms, presenting shorter deviants does not elicit MMN any more (Grimm, Roeber, Trujillo-Barreto, & Schröger, 2006), showing that the initial part of the stimulus is essential to form an object representation. To adapt this methodology to rats will certainly necessitate optimizing the procedure, by narrowing the ITI range and mixing the various deviants during the same sound sequence (Pakarinen, Takegata, Rinne, Huotilainen, & Näätänen, 2007) to keep such an experiment short enough.

The possibility to establish the time discrimination curve without the drawbacks exposed in the Introduction opens interesting perspectives from an evolutionary point of view. Indeed, it

has been shown that timing performance (as assessed by the variability of timing behavior) improves across species (Lejeune & Wearden, 1991), since the variability in time estimation decreases as we go up in the phylogenetic scale. However, between-species and between-ages differences might also reflect, at least partly, sub-optimal non-timing processes, among which attention is a critical factor, precluding onto- and phylogenetic comparisons. The approach we introduce here circumvents these limitations, since MMN mainly involves pre-attentive processes, and implies only very elementary sensory memory processes (see Schröger, 2007, for a review). As a MMN has been observed in several species, including monkeys (Javitt, Schroeder, Steinschneider, Arezzo, & Vaughan, 1992), rats (Eriksson & Villa, 2005; Ruusuvirta et al., 1998), guinea pigs (McGee et al., 2001), mouse (Umbricht et al., 2005), etc., (see Näätänen et al., 2007), those first data pave the way for the possibility to estimate, under the very same conditions and without attentional confounds, time discrimination threshold in different species to clarify the phylogenetic evolution of time processing. The same logic can be applied to compare the ontogenetic evolution of time estimation (Brannon et al., 2004). Indeed, timing performance has been shown to be U-shaped along the ontogenetic scale, with human adults being more accurate than children (Droit-Volet, 2002) and elderly people (McCormack, Brown, & Maylor, 2002). As for phylogeny, attentional confounds may contribute to this pattern of result (Vanneste & Pouthas, 1999).

The fact that a time discrimination threshold can be obtained in rat without behavioral tasks also opens interesting perspectives for other manipulations. For example, dopamine (DA) has been argued to be one of the central neurotransmitters involved in time estimation (Meck, 1996). However, DA manipulations not only impact timing processes, but also other cognitive operations (Nieoullon, 2002). Thus, as for phylogenetic comparisons, the results showing degradation of timing performance after DA depletion incur the risk of non-timing deficits degrading timing performance. Testing the impact of DA depletion on the discrimination threshold as assessed by the MMN would provide essential additional information about the pharmacology of interval timing.

## REFERENCES

- Brannon, E. M., Libertus, M., Meck, W. H., & Woldorff, M. G. (2008). ERP correlates of time perception in infants and adults: Weber's law holds. *Journal of Cognitive Neuroscience*, 20, 193–203.
- Brannon, E. M., Roussel, L. W., Meck, W. H., & Woldorff, M. G. (2004). Timing in the baby brain. *Cognitive Brain Research*, 21, 227–233.
- Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Review Neuroscience*, 6, 755–765.
- Burle, B., & Casini, L. (2001). Dissociation between activation and attention effect in time estimation: Implications for internal clock models. *Journal of Experimental Psychology: Human Perception and Performance*, 27(1), 195–205.
- Buzsáky, G. (2006). *Rhythms of the brain*. New York: Oxford University Press.
- Casini, L., & Macar, F. (1997). Effects of attention manipulation on perceived duration and intensity in the visual modality. *Memory & Cognition*, 25, 812–818.
- Coull, J. T., Vidal, F., Nazarian, B., & Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science*, 303, 1506–1508.
- de Lafuente, V., & Romo, R. (2005). Neuronal correlates of subjective sensory experience. *Nature Neuroscience*, 8, 1698–1703.
- Droit-Volet, S. (2002). Scalar timing in temporal generalization in children with short and long stimulus durations. *Quarterly Journal of Experimental Psychology A*, 55, 1193–1209.
- Eriksson, J., & Villa, A. E. P. (2005). Event-related potentials in an auditory oddball situation in the rat. *BioSystems*, 79, 207–212.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, 84, 279–325.
- Grimm, S., Roeber, U., Trujillo-Barreto, N. J., & Schröger, E. (2006). Mechanisms for detecting auditory temporal and spectral deviations operate over similar time windows but are divided differently between the two hemispheres. *NeuroImage*, 32, 275–282.
- Horváth, J., Cziger, I., Jacobsen, T., Maess, B., Schröger, E., & Winkler, I. (2008). MMN or no MMN: No magnitude of deviance effect on the MMN amplitude. *Psychophysiology*, 45, 60–69.
- Jacobsen, T., & Schröger, E. (2003). Measuring duration mismatch negativity. *Clinical Neurophysiology*, 114, 1133–1143.
- Jaramillo, M., Paavilainen, P., & Näätänen, R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters*, 290, 101–104.
- Javitt, D. C., Schroeder, C. E., Steinschneider, M., Arezzo, J. C., & Vaughan, H. G. J. (1992). Demonstration of mismatch negativity in

- the monkey. *Electroencephalography and Clinical Neurophysiology*, 83, 87–90.
- Knight, R. T., Brailowsky, S., Scabini, D., & Simpson, G. V. (1985). Surface auditory evoked potentials in the unrestrained rat: Component definition. *Electroencephalography and Clinical Neurophysiology*, 61, 430–439.
- Kujala, T., Kallio, M., Tervaniemi, R., & Näätänen, R. (2001). The mismatch negativity as an index of temporal processing in audition. *Clinical Neurophysiology*, 112, 1712–1719.
- Lee, K.-H., Egleston, P. N., Brown, W. H., Gregory, A. N., Barker, A. T., & Woodruff, P. W. R. (2008). The role of the cerebellum in sub-second time perception: Evidence from repetitive transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, 19, 147–157.
- Lejeune, H., & Jasselette, P. (1986). Accurate DRL performance in the pigeon: Comparison between perching and treadle pressing. *Animal Learning and Behavior*, 14, 205–211.
- Lejeune, H., & Wearden, J. H. (1991). The comparative psychology of fixed-interval responding: Some quantitative analyses. *Learning and Motivation*, 22, 84–111.
- Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognitive Brain Research*, 21, 139–170.
- McCormack, T., Brown, G. D. A., & Maylor, E. A. (2002). Effects of aging on absolute identification of duration. *Psychology and Aging*, 17, 363–378.
- McGee, T. J., King, C., Tremblay, K., Nicol, T. G., Cunningham, J., & Kraus, N. (2001). Long-term habituation of the speech-elicited mismatch negativity. *Psychophysiology*, 38, 653–658.
- Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cognitive Brain Research*, 3, 227–242.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118, 2544–2590.
- Nieoullon, A. (2002). Dopamine and the regulation of cognition and attention. *Progress in Neurobiology*, 67, 53–83.
- Pakarinen, S., Takegata, R., Rinne, T., Huottilainen, M., & Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clinical Neurophysiology*, 118, 177–185.
- Paxinos, G., & Watson, C. (1986). *The rat brain in stereotaxic coordinates*. New York: Academic Press.
- Rammsayer, T. H., Hennig, J., Haag, A., & Lange, N. (2001). Effects of noradrenergic activity on temporal information processing in humans. *Quarterly Journal of Experimental Psychology B*, 54, 247–258.
- Rammsayer, T. H., & Ulrich, R. (2005). No evidence for qualitative differences in the processing of short and long temporal intervals. *Acta Psychologica*, 120, 141–171.
- Ruusuvirta, T., Penttonen, M., & Korhonen, T. (1998). Auditory cortical event-related potentials to pitch deviances in rats. *Neuroscience Letters*, 248, 45–48.
- Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 62, 437–448.
- Schröger, E. (2007). Mismatch negativity: A microphone into auditory memory. *Journal of Psychophysiology*, 21, 138–146.
- Sussman, E. S. (2007). A new view on the MMN and attention debate. *Journal of Psychophysiology*, 21, 164–175.
- Thomas, E. A. C., & Weaver, W. B. (1975). Cognitive processing and time perception. *Perception & Psychophysics*, 17, 363–367.
- Umbricht, D., Vyssotki, D., Latanov, A., Nitsch, R., & Lipp, H.-P. (2005). Deviance related electrophysiological activity in mice: Is there a mismatch negativity in mice? *Clinical Neurophysiology*, 116, 353–363.
- Vanneste, S., & Pouthas, V. (1999). Timing in aging: The role of attention. *Experimental Aging Research*, 25, 49–67.
- Wearden, J. (2003). Applying the scalar timing model to human time psychology: Progress and challenges. In H. Helfrich (Ed.), *Time and mind II* (pp. 21–39). Göttingen, Germany: Hogrefe & Huber.

(RECEIVED July 28, 2008; ACCEPTED November 6, 2008)

# Spatio-temporal dynamics of information processing in the Brain: Recent advances, current limitations and future challenges

**Boris Burle<sup>a</sup>, Clémence Roger<sup>a</sup>, Franck Vidal<sup>a</sup>, Thierry Hasbroucq<sup>a</sup>**

*<sup>a</sup>Laboratoire de Neurobiologie de la Cognition,  
Aix-Marseille Université, CNRS, Marseille, France*

Correspondence: B Burle, LNC-Pôle 3C, Université de Provence, Case C, 3 Place Victor Hugo, 13331 Marseille, cedex 3, France. E-mail: [boris.burle@univ-provence.fr](mailto:boris.burle@univ-provence.fr), phone +33 4 88 57 68 79, fax +33 4 88 57 68 79

---

**Abstract.** Understanding how the brain transforms sensory information into adapted motor behavior necessitates to track the flow of information in the brain space. One question of great importance is to what extent the various required cognitive operations overlap in time (can a response begin to be prepared before the end of stimulus evaluation?). Symmetrically, it is essential to understand what is the degree of localization of the elementary cognitive operations (Are the motor areas purely motor, or do they intervene also in sensory processing?). After a brief statement of the current theoretical questions, we will present some recent data regarding these issues. The general logic followed is the track the information flow backward, starting from the response up to the stimulus. We will then present some technical limitation hampering more precise investigations and conclude on the challenges for the next few years for real advancement on those topics

**Keywords:** Information processing, motor processing, EEG, Spatio-temporal dynamics

---

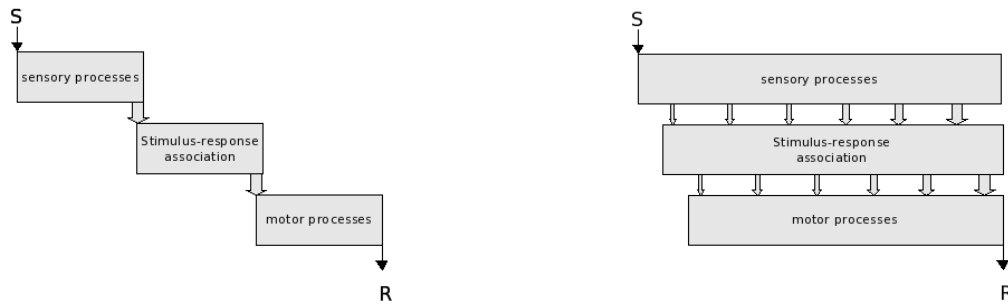
## 1. Introduction

We continuously receive sensory information from the environment. For our behaviors to be adapted to this environment, this sensory information has to be transformed into a motor code, through which our interactions with the environment are expressed. The question as to how the sensory codes are transformed into motor ones is one of the current challenges in cognitive neuroscience. One way to address this general question is to study the spatio-temporal dynamics of information processing in the brain, that is how the processing operation are organized in time and in space (of the brain areas). Both temporal and spatial dynamics can be considered along a continuum depending on the amount of overlap (both temporal and spatial) of the processing operations. We shall first present these two continuums. We will then present some how one can address those issues based on recent data and then raise some technical problems impeding further progresses on these issues.

### 1.1. Temporal overlap

A minimal decomposition of information processing requires three types of operations: sensory processes, stimulus-response association and motor processes. The temporal dynamic of such processes can be described on a continuum from purely serial processes to parallel ones. In the serial extreme, process  $n+1$  (for example stimulus-response association) cannot begin until process  $n$  has finished. Symmetrically, completely continuous models assert that as soon as process  $n$  receives any piece of evidence, it transmits continuously this information to process  $n+1$  (see [Sternberg, 2001] for an overview). Along this continuum, a large palette of architecture is conceivable. It can even be conceived that the flow of information is not purely feed-forward, but that feedback from level  $n+1$  on level  $n$  may occur. One important question is where, and under which conditions, cognitive processes lie on this continuum. For years, those questions have been addressed thanks to behavioral techniques. However, since the critical variable is the transmission of information between levels, only studying the product of the chain does not seem appropriate and investigating brain processes implementing those cognitive operations seems necessary.

Although EEG and MEG seem *a priori* the methods of choice, as we shall see later, there are currently technical limitations that impede such investigations.



**Figure 1.** Schematic representation of the two extremes forms of information processing temporal organization: left, purely serial model; right, purely continuous model.

## 1.2. Spatial overlap

If one aims at understanding the dynamics of information processing, one needs to match neurophysiologic indices with cognitive operations. The way such a matching can be done largely depends on the conception one has of the so-called structure-to-function relationship. The history of neuroscience has oscillated between two extremes points of view: localization versus distribution of functions. At that stage, one clarification is in order since a common mistake pollutes this debate: the discussion of localization vs. distribution should be conceived only at the level of elementary functions. As elementary as it could be debating on whether a behavior is localized or distributed has no sense. Even the simplest behavior, requiring several cognitive operations, is necessarily distributed. The relevant question is whether the *elementary operations that compose* such a behavior are or not distributed over the brain. For example, a question like “does reaching a pen lying on a table require distributed areas?” is incorrectly posed, since such a *behavior* requires seeing the pen, planning a movement and executing it. An example of a more relevant question would be whether, in such a motor context, the visual processing of the pen also involved non *a priori* visual areas, like the motor ones [Requin and Riehle, 1995].



**Figure 2.** Schematic representation of pure localization of functions (left) and largely distributed functions. The black dots (●) represent visual neurons, whereas the white ones (○) represent motor neurons.

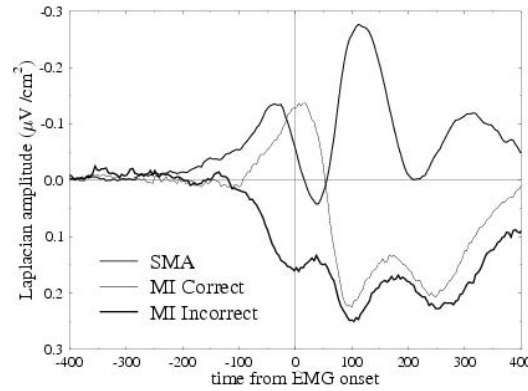
## 1.3. Orthogonality of these two continuums

For years, these two continuums have been confounded more or less implicitly on only one dimension, ranging from serial/localized to continuous/distributed. For example, “early” activation of the motor cortices has often been interpreted as reflecting continuous processes. However, the temporal and spatial dimensions *can* and *must* be dissociated. Indeed, operations may overlap in time (motor response could be prepared before stimulus evaluation is over) while these two operations remain perfectly separated in space (sensory processing limited to sensory areas, and motor processes to motor areas). Symmetrically, motor processes might well start after sensory processes have finished (no temporal overlap), but sensory and motor processes could partially be implemented in the *same* areas (spatial overlap).

Therefore, if one wants to use neurophysiological indices to track the temporal dynamic of information processing operations, it is necessary to also address their degree of spatial overlap.

## 2. Backwards analysis of the information processing flow.

For several years now, we have been studying the information processing flow in a backward fashion: starting from the response (*i.e.* the output of the system) to the stimulus. Such an approach has been motivated by the fact that studying the late stages of information processing makes it more likely to see the impact of previous stages on the current processing. In several studies, we first used electromyographic (EMG) activity that allowed us to evidence which experimental factors affect peripheral motor processes and which do not. We then went one-step further by analyzing information processing in the primary motor cortices (MI). In a between-response choice context, and by means of stimulation techniques (Hoffman reflex at the spinal level and transcranial magnetic stimulation – TMS), we have shown that just before the EMG onset the contralateral MI (involved in the incorrect response) is being *inhibited* [Burle et al., 2002; Hasbroucq et al., 2000] while the primary motor (MI) involved in the correct response is being activated. Such a activation/inhibition pattern is also observable with EEG [Tandonnet et al., 2006; Vidal et al., 2003], provided that appropriate deblurring methods (*e.g.* Surface Laplacian estimation) are used. Interestingly, the activities of the MIs are sensitive to some experimental manipulations (*e.g.* preparation [Tandonnet et al., 2006]), but others, more central, manipulations leave them unaffected. This differential sensitivity allows to investigate how and when information flow propagates to motor processes.



**Figure 3.** Time course of the activity recorded over the MI, and likely the SMA, around the response in a choice reaction time task. While the “correct” MI (*i.e.* contralateral to the correct response) is getting more active, the ipsilateral MI is inhibited. A fronto-central wave, likely originating from the SMA occurs about 50 ms before the MI. Adapted after [Vidal et al. 2003]

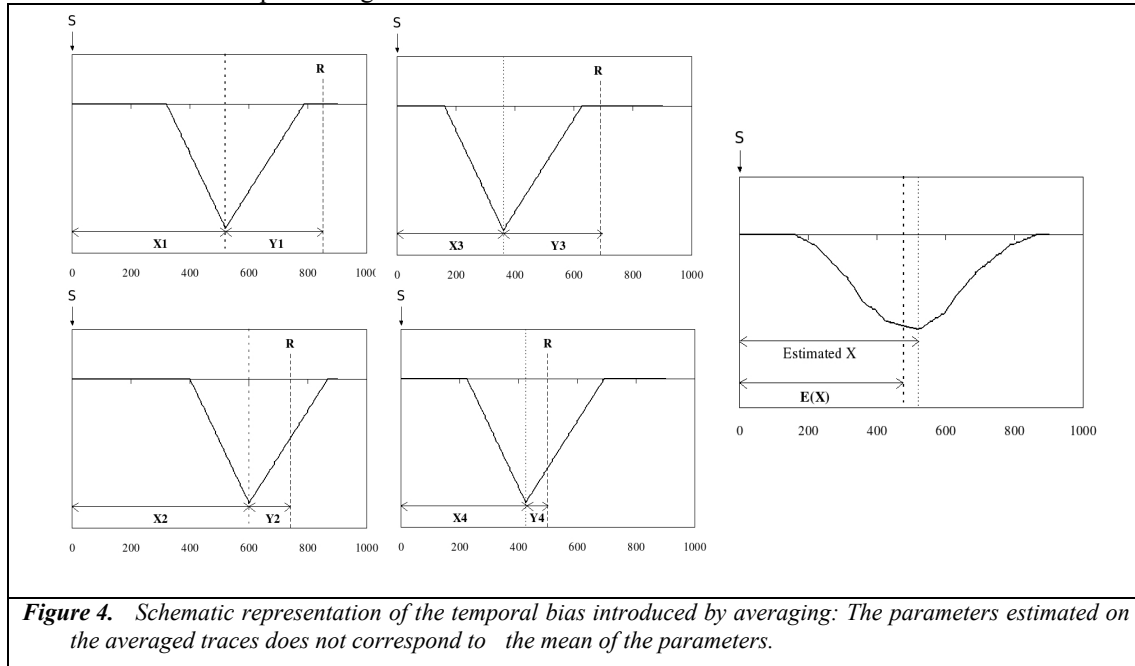
Laplacian estimation also revealed that premotor areas (like the Supplementary Motor Area – SMA) are activated slightly before the MI [Vidal et al., 2003]. Figure 3 shows an example of both the two MIs (electrodes C3 and C4) and the SMA (electrode FCz). In this particular case, the peak observed over the SMA occurred about 50 ms before MIs. This activity disappears when the subjects know in advance which response they will have to perform. Interestingly, the positivity observed ipsilateral to the correct response is also greatly reduced [Carbonnell et al., 2004], or even disappears, in this situation. This fronto-central activity might thus reflect either the inhibition of the not-to-be-given response, or the response-selection process.

## 3. Current technical limitations

### 3.1. The myth of the excellent temporal resolution of EEG/MEG: impact of averaging procedures

EEG and MEG are classically considered as methods of choice to study temporal dynamics, since their temporal resolution is excellent. However, this theoretical excellent resolution is largely degraded by the averaging process that is (currently) necessary to get an acceptable Signal/Noise ratio. Indeed, as already acknowledge more than 20 years ago by [Callaway et al., 1984] “The latency of the average is not the average of the latencies”. This is exemplified on Figure 4 that shows hypothetical single trials activities, along with their corresponding average. On the average trace, are reported both the true mean of the peaks latencies  $E(X)$  and the estimated mean value, corresponding to the peak of the average. As one can see, the estimated  $X$  clearly differs from the true  $E(X)$ : averaging induces a temporal bias, whose exact amplitude depends on the shape of the individual traces and on their

distribution latencies. As a matter, despite its name, the averaged trace does not reflect the central tendency of the individual traces. It should be better considered the convolution of the individual shape and a probability distribution of their latencies. Such a bias precludes precise investigations of the time course of information processing.



### 3.2. Improving the spatial resolution of EEG and MEG to secondarily improve their temporal resolution

Because of volume conduction, the spatial resolution of conventional EEG is very low. Importantly, as clearly pointed out by [Law et al., 1993], this bad spatial resolution also induces a degradation of the temporal resolution. Indeed, since volume conduction blurs the signal, if two electrodes are located above areas activated at slightly different moment, they will both record the same blurred signal, hindering the original temporal difference (see [Vidal et al., 2003] for an example). As indicated above, some methods exist to improve the spatial resolution in the captors space, like Surface Laplacian estimation. However, if Laplacian reduces this problem, it does not eliminate it. Hence, analyzing the signal in the source space seems desirable.

## 4. Challenges for the future

In order to improve your estimates of spatio-temporal dynamics of information processing, three complementary directions must be followed.

### 4.1. Estimating the signal of interest on a trial-by-trial basis

In order to precisely estimate the time course of brain activities, averaging should be avoided. This implies that we must estimate the “signal of interest” on a trial-by-trial basis. This implies either denoising techniques [Quian-Quiroga, 2000 ; Benkherraf et al., 2005] and/or parametric estimation of the noise [Bijma et al., 2005] and of the signal in parallel.

### 4.2. Estimating the activity of defined regions without contamination by surrounding regions

Although powerful source localization algorithms have been developed, their use to address information processing has been quite anecdotal so far. This might come from the fact that most of those algorithms are still heavy to use (MRI of the subjects, lengthy computations etc...), making psychologists and neuroscientist not really incline to use them. Developing tools that remain compatibles with the constraints of regular psychological and/or neurophysiological experiments (large number of subject, quite a lot of conditions etc...) is necessary.



### 4.3. Estimating the communication between areas

As indicated in the introduction, continuous and serial models differ in the way information is transmitted from one stage to the next. It is thus necessary to measure which and how brain areas communicate. Ideally, such between areas communication should be detected on a trial-by-trial basis, to avoid the bias evoked above and to allow a better correlation with behavior.

## 5. Conclusions

In this brief overview, we tried to present some of the current questions in the field of cognitive neuroscience, along with the general methodology to address these questions. We have also tried to show the current technical limitations that impede more precise investigation of those issues, together with some directions to solve those technical problems.

## Acknowledgements

The manuscript was partly supported by a “Neuroinformatique” grant from the CNRS to the first author.

## References

- Benkherraf M, Burle B, Allain S, Hasbroucq T, Vidal F. Individual Evoked Potential Extraction by Multiresolution Wavelets Decomposition. In proceedings of EUROCON 2005, The International Conference on Computer as a Tool, 1: 417-420, 2005.
- Bijma F, de Munck JC, Heethaar RM. The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker products. , *NeuroImage*, 27: 402-415, 2005.
- Burle B, Bonnet M, Vidal F, Possamaï CA, Hasbroucq T. A transcranial magnetic stimulation study of information processing in the motor cortex: relationship between the silent period and the reaction time delay. *Psychophysiology*, 39: 207-217, 2002.
- Callaway E, Halliday R, Naylor H, Thouvenin D. The latency of the average is not the average of the latencies. *Psychophysiology*, 21: 571, 1984.
- Carbonnell L, Hasbroucq T, Grapperon J, Vidal F. Response selection and motor areas: a behavioural and electrophysiological study, *Clinical Neurophysiology*, 115: 2164-2174, 2004.
- Hasbroucq T, Akamatsu M, Burle B, Bonnet M, Possamaï CA. Changes in spinal excitability during choice reaction time: The H-reflex as a probe of information transmission. *Psychophysiology*, 37:385-393, 2000.
- Law SK, Rohrbaugh JW, Adams CM, Eckardt MJ. Improving spatial and temporal resolution in evoked EEG responses using surface Laplacians. *Electroencephalography and Clinical Neurophysiology*, 88, 309-322, 1993.
- Quiñero Quiroga R. Obtaining single stimulus evoked potentials with wavelet denoising. *Physica D*, 145: 278-292, 2000.
- Requin J, Riehle A. Neural correlates of partial transmission of sensorimotor information in the cerebral cortex. *Acta Psychologica*, 90: 81-95, 1995.
- Sternberg S. Separate modifiability, mental modules, and the use of composite and pure measures to reveal them. *Acta Psychologica*, 106: 147-246, 2001.
- Tandonnet C, Burle B, Vidal F, Hasbroucq T. Knowing when to respond and the efficiency of the cortical motor command: A Laplacian ERP study. *Brain Research*, 1109: 158-163, 2006.
- Vidal F, Grapperon J, Bonnet M, Hasbroucq T. The nature of unilateral motor commands in between-hands choice tasks as revealed by surface Laplacian estimation. *Psychophysiology*, 40: 796-805, 2003.



## **Supervision de l'action et optimisation des comportements : Études électrophysiologiques et IRMf**

Les mécanismes cérébraux impliqués dans l'optimisation des comportements sont encore mal connus. Nous avons cherché à préciser leur fonctionnement à travers des études EEG et IRMf dans des tâches de temps de réaction manipulant la compatibilité Stimulus-Réponse.

Plusieurs arguments issus de méthodologies différentes nous ont permis de préciser le rôle de la Négativité d'Erreur ( $N_E$ ), une onde EEG initialement rapportée dans les erreurs, et qui atteint son maximum juste après la réponse. Initialement attribuée à la détection de l'erreur, cette activité, localisée dans la zone cingulaire rostrale, pourrait avoir pour rôle d'évaluer la performance en cours d'essai dans le but de rattraper l'erreur.

Nous nous sommes intéressés aux mécanismes impliqués de la sélection de la réponse. La N-40, enregistrée au-dessus de l'aire motrice supplémentaire, incarnerait l'association stimulus-réponse alors que les cortex moteurs seraient le lieu de la mise en place d'un mécanisme de prévention de l'erreur implémenté par une élévation du seuil de déclenchement des réponses en présence d'un risque d'erreur.

Enfin dans une étude IRMf, nous avons montré que les ajustements comportementaux suite à un essai incompatible prenaient la forme d'une augmentation du traitement perceptif des caractéristiques pertinentes de la tâche et d'une diminution du traitement des caractéristiques non pertinentes.

Nous avons précisé le rôle fonctionnel de la  $N_E$  ainsi que les opérations mises en jeu lors de la mise en place des ajustements comportementaux. Ces résultats suggèrent également l'existence de mécanismes capables de détecter en ligne un risque d'erreur dans le but de prévenir une erreur en cours d'essai.

### **Action monitoring and optimisation of performance : EEG and fMRI studies**

We studied the cerebral mechanisms involved in the optimisation of performance. In order to increase the knowledge of these mechanisms, we ran EEG and fMRI experiments in reaction time task in which we manipulated the stimulus-response compatibility.

Firstly, we investigated the role of the negativity of error ( $N_E$ ), an EEG component which has been discovered in errors and which reaches its maximum just after the response. The  $N_E$  was first interpreted as reflecting the error detection. We suggest that the  $N_E$ , localised in the rostral cingulate zone, is involved in the evaluation of the performance in order to correct erroneous responses.

Secondly, we were interested in mechanisms involved in response selection. The N-40, recorded over the supplementary motor area, is proposed to play a role in the response selection stage. Error prevention would take place at the primary motor cortices level by an enhancement of the response threshold when necessary.

Thirdly, we showed, in a fMRI study, that the behavioural adjustment after an incompatible trial is due to an increase of the processing of relevant feature of the stimulus and a decrease of the irrelevant one.

In this work, we clarified the role of the  $N_E$  and the processes which are responsible for the behavioural adjustments. These results suggest the existence of a mechanism able to detect on-line the error risk involved in preventing errors.